ARI Research Note 96-16

Decision Making Under Uncertainty and Time Stress

John Leddo, Mike O'Connor, Julie Doherty, and Terry Bresnick
Decision Science Consortium, Inc.

Fort Leavenworth Research Unit
Stanley M. Halpin, Chief

January 1996

United States Army
Research Institute for the Behavioral and Social Sciences

U.S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Director

Research accomplished under contract
for the Department of the Army

Decision Science Consortium, Inc.

Technical review by

James W. Lussier

NOTICES

DISTRIBUTION: This report has been cleared for release to the Defense Technical Information Center (DTIC) to comply with regulatory requirements. It has been given no primary distribution other than to DTIC and will be available only through DTIC or the National Technical Information Service (NTIS).

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The views, opinions, and findings in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

1. REPORT DATE 1996, January	2. REPORT TYPE Final	3. DATES COVERED (from. . . to) December 1988-December 1990

4. TITLE AND SUBTITLE Decision Making Under Uncertainty and Time Stress	5a. CONTRACT OR GRANT NUMBER MDA903-89-C-0031 5b. PROGRAM ELEMENT NUMBER 0605502A
6. AUTHOR(S) John Leddo, Mike O'Connor, Julie Doherty, and Terry Bresnick	5c. PROJECT NUMBER M770 5d. TASK NUMBER 1102 5e. WORK UNIT NUMBER C01
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Decision Science Consortium, Inc. 1895 Preston White Drive Suite 300 Reston, VA 22091	8. PERFORMING ORGANIZATION REPORT NUMBER 91-1
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences ATTN: PERI-RS 5001 Eisenhower Avenue Alexandria, VA 22333-5600	10. MONITOR ACRONYM ARI 11. MONITOR REPORT NUMBER Research Note 96-16

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

This is a Small Business Innovative Research Program, Phase II Report.

14. ABSTRACT *(Maximum 200 words)*:

Tactical decision making occurs in environments characterized by high uncertainty, high-valued outcomes, and time stress. The critical relation of decision outcomes to uncertain events pushes decision makers to resolution of uncertainty. Time stress, however, works in the opposite direction, precluding ideal planning efforts. It is thus crucial to systematically characterize the impacts of uncertainty and time stress on tactical decision making to facilitate the design of decision aids that will be effective in the environments described.

The present study employs a theoretical framework that extends Kahneman and Tversky's prospect theory (1979) to goal-directed decisions such as those involved in battlefield tactical decisions. The framework is used to provide a basis for characterizing the critical issues involved in such environments and to design decisions aiding concepts based on studies employing the framework.

Four studies were conducted. The first study provided participants with choices involving gains and losses in goal-directed tactical and nontactical situations. The purpose was to examine preferences among risky options and to validate the extended theory as an appropriate characterization of the tactical decision making process. The results strongly confirmed the theoretical (Continued)

15. SUBJECT TERMS

Decision making   Uncertainty   Time stress   Decision aiding

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number)
16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified	Unlimited	130	

14. ABSTRACT (Continued)

formulation. The second study involved a more rigorous test of the framework using a more realistic tactical problem that provided both for a test of the framework and development of concepts for aiding decisions under time stressed conditions. Tradeoffs between effort allocated to uncertainty reduction and option generation were examined in detail, and aid concepts were elicited. Study two results were used to design aiding concepts that were further refined and tested in study three to yield a final set of aiding concepts to be tested in study four.

Study four examined the decision time and decision quality impacts of aiding tactical decision makers through the use of more decision makers (group), decision aids, or both. The results indicated that aiding decision makers did not cause a significant increase in time taken to arrive at a course of action recommendation. For individual decision makers, providing an aid impacted decision quality through increased situation assessment accuracy. For group decision makers, the use of aids influenced course of action selection in a manner that was not attributable to situation assessment accuracy. Further, aided groups tended to recommend risky courses of action that were unlikely to succeed given the actual situation. Results suggest that studies of tactical decision making and planning must acknowledge the goal-directed nature of such endeavors and must accommodate this issue in aids developed. Further, designers of aids must carefully address the organizational issues of individuals versus groups using such aids and must be careful with respect to both individual and group tendencies toward biased decision making that may interact with usage of aids.

DECISION MAKING UNDER UNCERTAINTY AND TIME STRESS

CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

DECISION MAKING UNDER UNCERTAINTY AND TIME STRESS

STUDY 1

INTRODUCTION

Objective

The research presented in Study 1 tested parts of a theoretical framework for assessing and predicting judgment and choice behavior for decisions potentially involving both large gains and large losses. The research focused on decisions that are goal-related and provide accomplishment-related feedback. Of interest was determining how both different levels of potential gain or loss and stated outcome probabilities influenced the decision process. Of special interest was determining how time stress would interact with the decision process, particularly in cases where the decision maker has insufficient time to assess risks fully and determine whether options being considered meet stated objectives embodied in the goals.

It is argued that the results of the research have general utility in describing and predicting military decision making, but can also have implications for decision aiding where decisions are associated with potentially large gains and also great risks. The results of the present study were used to develop concepts for aiding decision makers under conditions of time stress and uncertainty and where large gains and losses are at stake. These concepts were examined in later experiments.

Prospect theory as explained by D. Kahneman and A. Tversky (1979) was used as a theoretical framework for the study. The concepts and results of studies related to this theory which was developed primarily for describing individual choice behavior under risky conditions were used as theoretical referents and data points with which to compare the military decision making process at the division planning level. By comparing the military decision making situation to that described by prospect theory, similarities and differences can be identified that will be valuable in identifying both inter-situational similarities in behavior as well as important differences that require further theoretical development and practical experimentation.

In this section, a general framework for goal-based decisions under uncertainty and time stress is outlined, and several testable hypotheses are described. Prior to discussing this framework, it is useful to introduce some terms and concepts that will be used.

Historical Perspective

These terms and concepts are those generally used in decision theory and thus have been around for over fifty years. In the late 1940's, an attempt was made to use the prescriptive theories of economics and statistics as a descriptive theory for individual choice behavior under uncertainty. Put very simply, any individual choice could be presented as a choice among two gambles:

$$\text{gamble 1} = g_1 = (a, P_1, b)$$

$$\text{gamble 2} = g_2 = (c, P_2, d)$$

Gamble 1, g_1 thus involves a lottery in which the outcome a (which could be a prize, an amount of money, or even event) occurs with probability P_1, and otherwise the outcome b occurs with probability $(1 - P_1)$. For g_2, outcome c occurs with probability P_2, and otherwise outcome d occurs with probability $(1 - P_2)$.

Economic theory prescribes that in choosing between the options g_1 and g_2, a rational, economic person should compute the expected value (EV) of each option and choose that option with the higher EV:

$$EV(g_1) = P_1 a + (1 - P_1)b$$

$$EV(g_2) = P_2 c + (1 - P_2)d$$

Descriptive experiments quickly indicated that people clearly did not do what economic theory prescribed--for example, consider:

$$g_1 = (\$10,\ .5,\ -\$5) \quad EV(g_1) = +\$2.50$$

$$g_2 = (\$1{,}000,\ .5,\ -\$950.00) \quad EV(g_2) = +\$25.00$$

Most people faced with a choice between g_1 and g_2 would choose g_1 even though the expected value of g_2 is ten times as much. Further economic theory prescribes that one should be willing to pay up to \$25.00 (say \$20.00) for the chance to play g_2. (\$25.00 is called the maximum buying price of g_2.) The theory also indicates that anyone having the gamble g_2 should not be willing to sell it for less than \$25.00 (The minimum selling price of g_2)

In truth, many people having g_2 would be quite willing to give it away or possibly even pay someone to take it! This is because of a phenomenon known as risk aversion. People are typically risk averse and thus buy life insurance even though it has a very negative expected value.

The concept of a subjective transformation on value (which denotes "objective" value, e.g. dollars) to yield the subjective value or utility of an outcome was introduced early on to explain such phenomenon as risk aversion. Consider:

$$g_1 = (\$10,\ .5,\ -\$10.00) \text{ and}$$

$$g_2 = (\$1{,}000,\ .5,\ -\$1{,}000)$$

These two gambles have the same expected value of zero, but most would prefer g_1 to g_2 if forced to choose. The concept of a marginally decreasing utility for money so that each incremental unit of value has slightly less subjective value than its predecessor is represented by a concave utility curve. A utility curve for losses such that each incremental unit of loss in value is slightly worse than its predecessor is represented by a convex utility curve for losses. Taken together as an individual's utility function over value, these curves predict choices such as g_1 over g_2. (Gambles g_1 and g_2 also have different variances, and variance in outcomes has also been

equated with risk.) Friedman and Savage (1948) described the utility curve of an individual that would both gamble (for relatively small amounts) and also buy insurance (against catastrophic losses).

Using such utility curves, rational, economic man was assumed to maximize expected utility (EU). For gambles g_1 and g_2, the expected utility of each gamble would be less than zero, the expected value of each. Further, assuming the utility of zero to be zero, the expected utilities of each of these gambles is less than the utility of its expected value (zero). If we denote "zero" value as "zero incremental wealth over the status quo level of wealth", than the utility curve described earlier that is concave for gains, convex for losses, would predict a choice of the sure thing status quo over either of the gambles g_1 or g_2.

Further theoretical and experimental work showed that the EU theory fails to describe individual choice behavior in many cases, and a subjective transformation over probability was introduced. The experimental finding that subjects overestimated very low probabilities of large negative losses and underestimated large probabilities supported the notion of a subjective transformation S on the probability function so that the gamble g_1 (a P b) has as its subjectively expected utility (SEU)...

$$SEU(g_1) = S(P)U(a) + S(1-P)U(b)$$

This theory, though mathematically quite general, still assumes that the S and U functions are assigned independently, whereas much of the experimental work produced dependencies between the values and probabilities -- e.g., probabilities associated with high gains or extremely likely outcomes were underestimated (where objective data existed), whereas probabilities of extremely unlikely or negative outcomes (e.g., cancer) were overestimated or assessed. Thus, experimental work progressed in questioning this economic theory as descriptive of decision making and thus promising in helping to predict behavior and possibly improve or aid it.

The utility theory thus far described views departures from optimal or normative economic behavior as being due to subjective transformations on the information at hand, i.e., values and probabilities. Around 1957, beginning with the work of Simon (Simon, H., 1957), a different view was promulgated, that of man as a semi-rational (some said irrational), certainly non-normative, limited information processor who had limited storage capacity, limited information processing capabilities, and limited ability to handle complexity. This man thus mis-perceived the "objective" data and applied often simple decision rules or "heuristics" in decision making. Kahneman and Tversky conducted a series of experiments predicated on this viewpoint (See Tversky, A. and Kahneman, D., 1974 and Kahneman et. al. (Eds), 1982) and revealed numerous departures from normality predicted from this viewpoint. Thus began the integration of economic theory and psychological (cognitive) theory in explaining choices under uncertainty and risk. (A distinction is sometimes made between risky decision making where outcomes have defined probabilities on a well understood, event mechanism generating them and choice under uncertainty where the probability is ill-defined. For the purposes here, the two will be equated).

Around this time, a different movement also occurred in organizational theory in which the behavior of organizations began to be described as if the

organization were an entity that behaved somewhat in the manner described by Simon and also Kahneman and Tversky (though they in fact came later). Thus, Cyert and March (1963) describe the behavior of the firm in making choices and exerting control as:

- having multiple, changing, accept ie level goals.
- as considering alternatives in an approximately sequential manner.
- as seeking to avoid uncertainty through the use of regular procedures (e.g., military doctrine) and reacting to feedback rather than forecasting the environment.
- using standard operating procedures and rules of thumb to make and implement choices.

Thus, the organization behaves somewhat like the decision maker postulated by Simon.

The details of these two vast literatures will not be provided here. However, the similarities should be fairly clearer, and the relevance to military decision making should also be clear. Military decision makers are members of a very large goal-directed organization that must make high stakes decisions in environments of high uncertainty, risk, and stresses of all kinds. Using these different frameworks as a comparison reference for studying such military decision making should prove highly productive. Continuing with the decision theory approach as modified by Kahneman and Tversky, prospect theory is of interest.

Prospect Theory

One of the most prominent and influential theories of decision making and choice behavior is Kahneman and Tversky's prospect theory (1979), which evolved from attempts to understand the place of expectation based theories in describing individual decision making. prospect theory proposes a value function which relates actual outcome value to subjective utility and a decision weighting function which translates the stated probability of an outcome to a subjective weight that the stated probability carries in assessing the attractiveness of that outcome.

The value function has two distinct properties: a) it is concave for gains and convex for losses so that for example, the difference in utility between 1,100 and 1,200 dollars is not the same as the difference in utility between 100 and 200 dollars, and; b) the function for losses is steeper than the function for gains so that a given amount of loss is more aversive than the same amount of gain is attractive. The decision weighting function has the properties that very small probabilities are overweighed, while moderate and large probabilities are underweighted. Figures 1 and 2 show prospect theory's value function and decision weighting function, respectively.

The properties of these two functions in describing individual decisions predict that: people are risk averse when it comes to gains and risk seeking when it comes to losses. For example, most people would rather have a sure $500 than a 50% chance to win $1,000 (in fact, they would even take a sure $400 to a 50% chance to win $1,000) but would rather have a 50-50 chance at losing either $1,000 or nothing than a sure loss of $500 (Kahneman and Tversky, 1979).

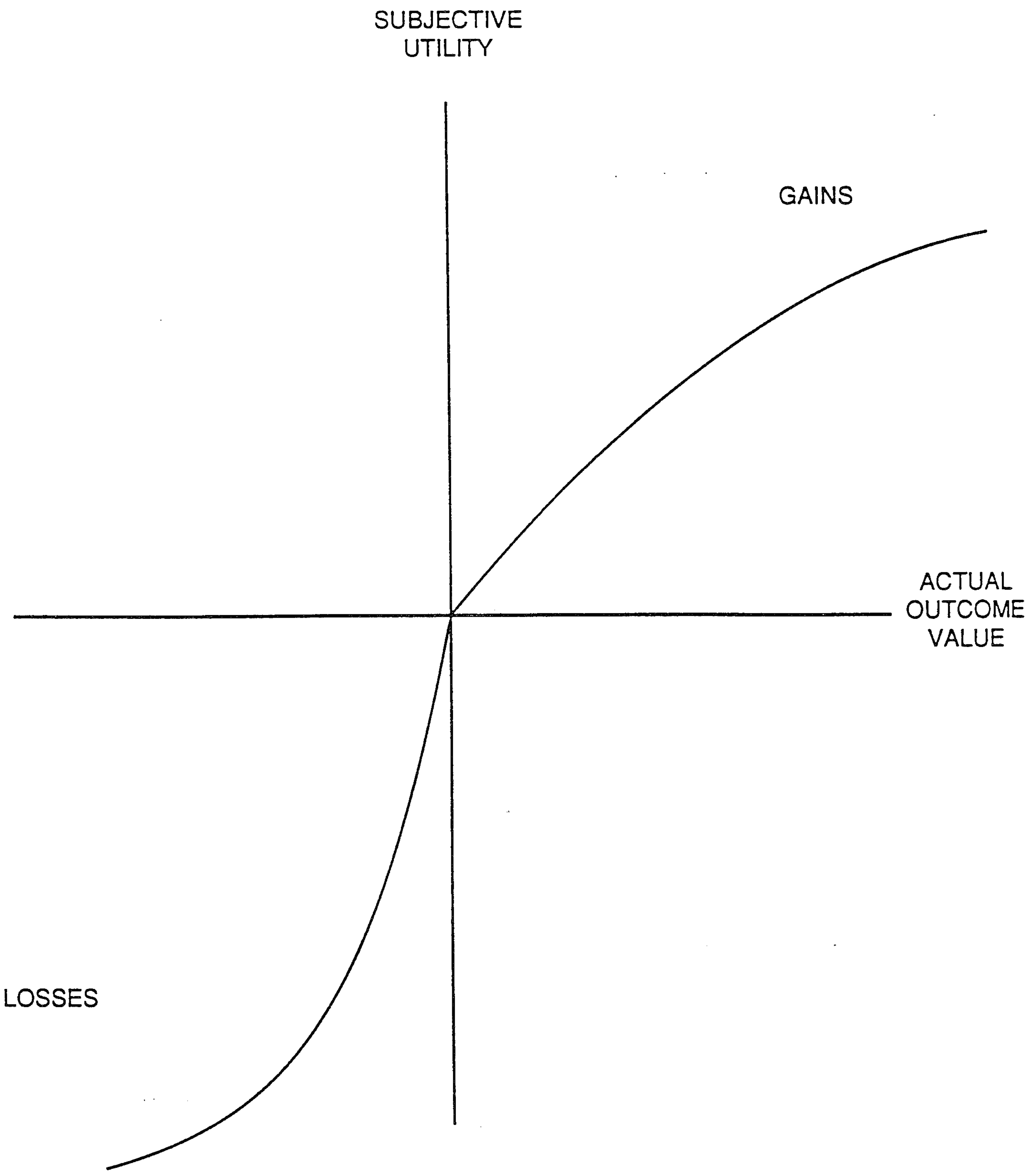


FIGURE 1. Kahneman and Tversky's (1979) value function showing relationship between actual outcome value and subjective utility.

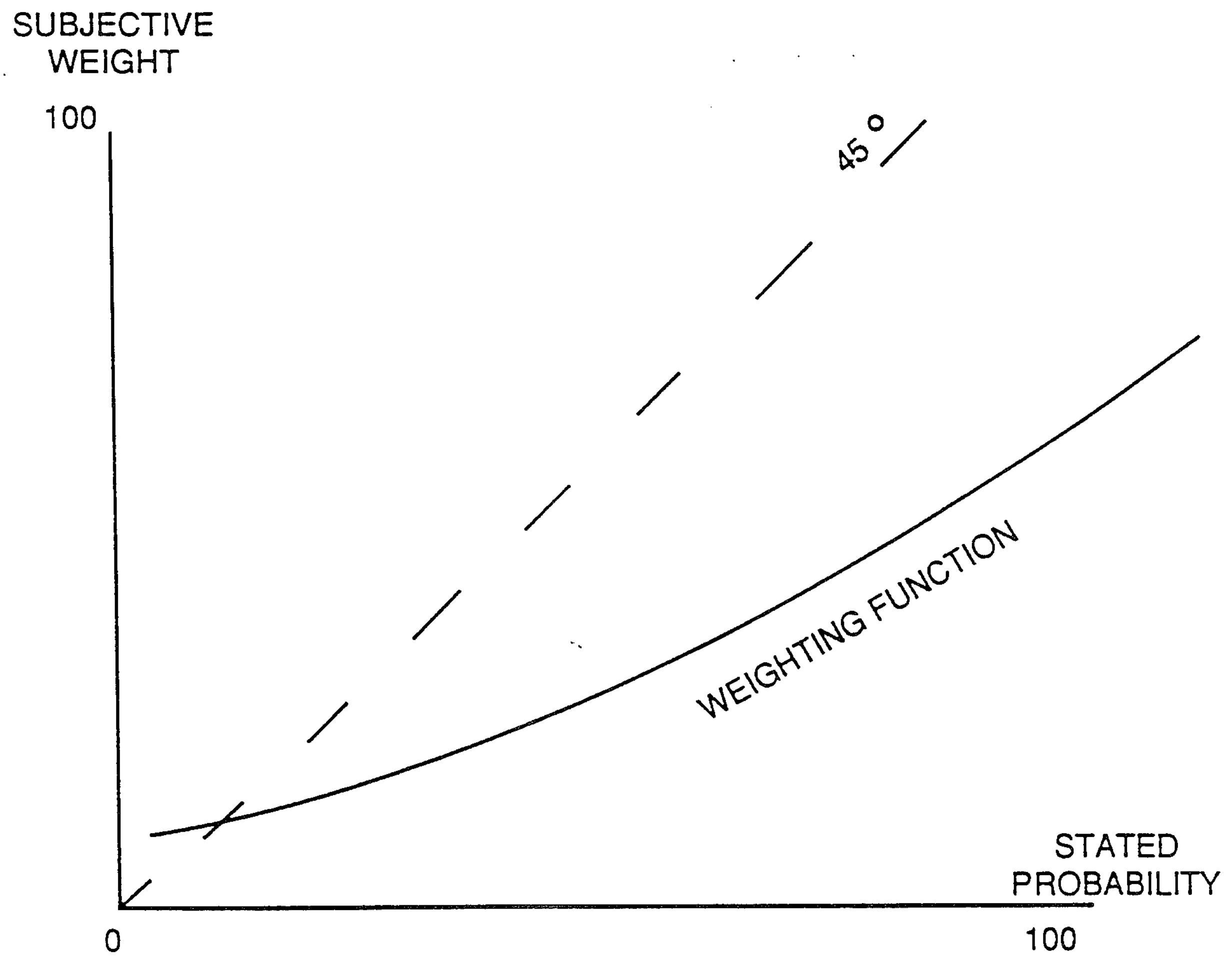


FIGURE 2. Kahneman and Tversky's (1979) decison weighting function showing relationship between stated probability of outcome and subjective probability.

One implication of prospect theory that is both interesting and troublesome to decision making theorists is that presenting (framing) the same pair of choices either in terms of their relative gains or their relative losses can lead the same person to pick opposite choices. In a classic example, (Tversky and Kahneman, 1981) people were given a scenario which involved trying to protect 600 potential victims of a disease through one of two programs. The effectiveness of the two programs was described either in terms of how many people would be saved or how many people would die. Specifically, in the save version, program A would save 200 lives whereas program B had a one-third chance of saving all 600 people and a two-thirds chance of saving no one. In the mortality version, program A would result in 400 deaths whereas program B would have a one-third chance of no deaths and a two-third chance of 600 deaths. Although the two versions describe the programs' effectiveness equally, when presenting the programs in terms of lives saved, 72% of the subjects surveyed preferred taking the sure saving of 200 lives to gambling on the lives of all 600 people, whereas presenting the programs in terms of lives lost, 78% of the subjects surveyed preferred risking all 600 lives to save potentially all of them as opposed to allowing 400 people to die for sure.

A very important theoretical issue that becomes a practical issue in using prospect theory (or any utility formulation) to describe military decisions concerns the maintaining of gains and losses. It is assumed that the decision maker considers options with respect to some reference point, say his status quo, and options are characterized as gains or losses from that status quo. Thus, in terms of preference orderings, given a series of gambles $[G_i]$ involving gains and losses, several important issues must be clarified. If the objective gain is g, (this implies g be measured on some "objective continuum) will obtain with "objective" or "stated" probability P, this means that the probability of the outcome that yields the gain g is well understood, possibly universally agreed upon as P, and the same holds true for the losses. Thus, the decision maker knows the objective probabilities. The utility (U) and probability (S) functions then represent functions which can be used to describe the results of the subjective transformation made by the decision maker on the gamble inputs (g, P, l) so that the subjective worth of the gamble $G_i = (g_i\ P_i\ l_i)$ is $U(G_i) = S(P_i)\ U(g_i) + S(1 - P_i)U(l_i)$ which is the earlier discussed subjectively expected utility theory. Prospect theory has added some very important psychological aspects to this theory by introducing gains, for the decision maker in choosing among the $\{G_i\}$ looks at each G_i from his status quo using the utility and probability functions described.

This is not a single utility function defined over some defined, invariant continuum, say total wealth. (Note, the continuum could be a single vector that is a linear combination of several others thus accommodating multi-dimensional value problems. It is required that the continuum is understood to the decision maker in some way and that the utility function is defined over this meaningful continuum.) Rather the utility function is defined over losses and gains, so that when the total wealth of an individual changes, his status quo changes, and options are considered from the new reference point which becomes the new origin for the same function.

Extension of Prospect Theory

Kahneman and Tversky have done an excellent job in explaining cases where people would be risk averse versus risk seeking, and have shown how framing choices can lead to one tendency or the other. The present study

extends the framework of prospect theory to the military planning and decision making context. A goal is to use this and related frameworks to better characterize military decisions. Specifically, prospect theory is extended along the following two dimensions:

1) Location of outcomes with respect to goal states. Most of Kahneman and Tversky's examples involve cases where people are presented with choices whose outcomes are not stated with respect to any goal (e.g. they would win or lose money without any reference to how much money they need). It is argued that many real life choices involve courses of action whose outcome are initially related to the goals of the person making that choice. This is especially true in the military decision making context where course of action generation is driven by unit missions or goals.

2) The relationship of the value function to the decision weighting function. Prospect theory assumes that the value function and the decision weight function are independent, e.g. a 50% chance of achieving a desired outcome carries the same weight regardless of whether the outcome is $1 or $1,000. More recently, Einhorn and Hogarth (1984) suggest that people's subjective probabilities may depend on the level of the outcome, although they do not fully describe what that relationship will be.

Having outlined the assumptions underlying prospect theory and the proposed extensions, hypotheses are discussed regarding the prospect theory to military decision making problems.

Prospect theory is particularly relevant to military decision making in that decision makers are goal (mission) oriented and must think in terms of gains and losses from their current positions, the status quo. However, there are several crucial differences between the choice context often discussed by Kahneman and Tversky and the military decision making situation.

In making plans to accomplish a mission, a military decision maker is constrained to a degree by guidance from higher authorities that sets political constraints, etc. Also the higher authority provides a mission for the military planner, and his job is to achieve the mission goal. Goal achievement thus is a threshold type of situation that is very likely to lead to behaviors that might be best characterized by discontinuous utility functions.

A second aspect of military decisions involves the usually high stakes. The military commander operates conditional on given doctrine and precedents. He also has guidance with respect to acceptable or expected casualty levels that are related to mission achievement. Failure to achieve the mission is his primary fear. A secondary fear or point to avoid becomes a level of catastrophic loss with respect to personnel and material. Given a mission to take terrain point X by time T and hold point X for at least 24 hours, it is quite useless to suffer losses in personnel and arms in taking the point such that it can only be held for a few moments. Thus if mission achievement maps into the outcome of taking terrain and holding it for a period of time, there are correlated measures of success such as movement at a certain rate toward X, maintaining firepower, minimizing casualties.

Mission accomplishment may not be all or none. It may be that taking a point quite close to X is acceptable. On the other hand, this may not be the case. If the purpose of taking point X is to engage the enemy and hold him in

place, the attempt to take X and associated enemy engagement may yield partial mission success. However, it is quite likely that for the military commander, the smooth curve in Figure 1 represented by prospect theory will not be applicable. This discussion yields several important points.

Unlike the decision maker in prospect theory choosing among lotteries or the like from his status quo reference point, the military decision maker considers options with respect to at least one and likely two reference points. These we will label the Aspiration level and the Avoidance Level as illustrated in Figure 3.

Consider mission achievement to be associated with the Aspiration level. Thus the single continuum illustrated in Figure 3 represents the actual outcome associated with mission achievement (note that this could actually be a point in multi-dimensional space characterized by acceptable levels along several dimensions). For example, the outcome continuum could be similar to the type of continuum in a c"critical incidents" type of scale. Mission success means taking the point X in time with sufficient remaining combat capability to hold it for the required time.

The status quo or current level may be another point along the continuum characterized by current disposition including location, strength, etc. A third point indicated on the continuum is labelled the "Avoidance level". Such a level could represent being driven back to a point worse than the current one with extremely high loss of personnel, weapons, and provisions. Thus it is not simply mission failure. In fact, given the mission goal, it is more like total defeat.

Considering these outcomes as points on a continuum is, of course a simplified representation, but it serves well to illustrate the concepts. Extension to spaces of higher dimensionality does not change the basic concepts.

A second difference between the prospect theory situation and the military context is that the military decision maker does not consider gains and losses with respect to the status quo. He has changing reference points, one being his Aspiration level -- mission success. A second point is his Avoidance level, total defeat. (Others may be relevant but these are the most salient.) He has no interest in maintaining the status quo or possibly, ever seeing it again. It is, however, a reference point for planning.

A third crucial difference between the military decision context and the prospect theory context concerns the independence of the utility function and the value function earlier discussed. In the military decision, the decision maker creates the options and he also creates the probabilities through course of action development (the g and p in (g, p, l) earlier discussed). Also, by creating his plan to achieve his mission, he also creates the potential loss that he might sustain and the associated probability of sustaining it.

One can see that characterizing the military decision problem as a two outcome gamble is far too simple. However, the concepts are still very valuable as a framework from which to consider military decision making. Military decisions are indeed choices among risks, but they are not games against an indifferent Mother Nature. By the wargaming process that goes on in planning, the second and third order dependencies in actions can be

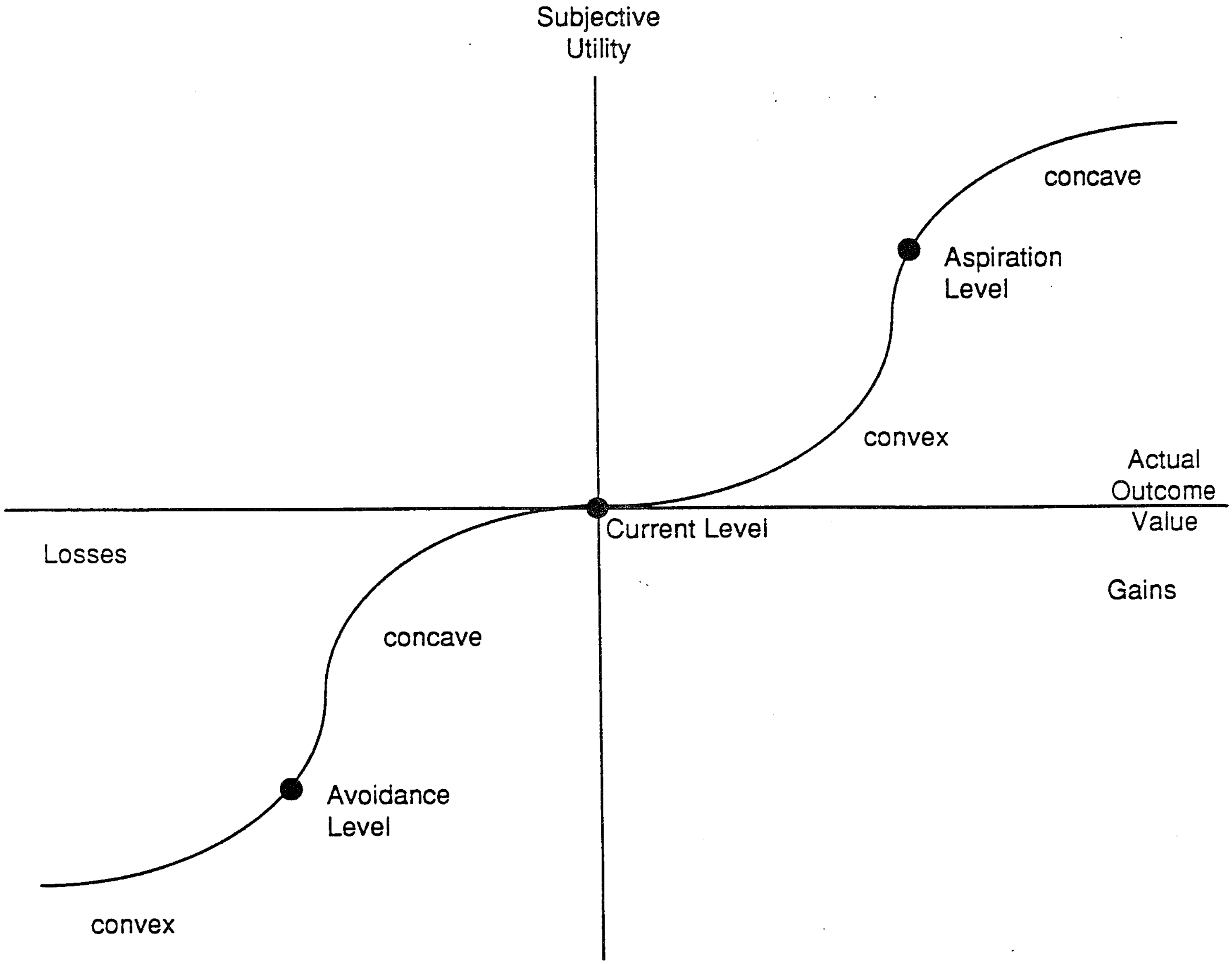


FIGURE 3. A value function showing aspiration and avoidance levels as reference points.

accommodated, and the commander's decision can still be characterized as choices among multi-outcome gambles.

The major point is that the probability of loss is correlated with the probability of gain. By attempting to achieve the mission, the commander increases his probability of sustaining severe losses. In most cases, the probabilities of mission achievement and total defeat will be related due to the inherent nature of military combat.

A fourth issue that arises is the discontinuous nature of military decision making. Decisions are conditional. No loss of life is ever acceptable. However, conditional on the requirement to achieve the mission, certain casualty levels are to be expected. Casualty levels, though important, are first considered with respect to mission success. The decision rule is discontinuous. The commander must consider his decisions almost sequentially and his preference ordering over outcomes is like the "lexicographic" ordering in which all outcomes involving mission success are favored over all other outcomes. Conditional on failure to achieve the mission, the Avoidance level becomes reference point number two, and all outcomes above the Avoidance level are preferred to any below it. Thus, casualties could then become the prime determinant of choices.

It is recognized that actual military decisions, do, in fact, involve tradeoffs with respect to success and casualties, but the point is that the commander is required to have "lexicographic" like decision functions, (named after the lexicographer who orders first on the basis of the first letter a, then the second letter b, etc. This is a discontinuous rule. A's come first no matter what.) He does not face a simple "choice among gambles" type of situation. Yet he must make choices among complex options with uncertain outcomes and high status.

<u>Framing effects and attitudes toward risk.</u> As discussed earlier, Tversky and Kahneman (1981) showed that people will often make different choices regarding the same set of options depending upon whether they view those choices in terms of gains or losses. It is argued that a similar effect will occur depending upon whether they are view those choices in terms of outcomes above or below the aspiration and avoidance levels. In fact, it is argued that people will use the aspiration and avoidance levels to frame decisions spontaneously. Figure 3 presents a revised value function which is in essence the prospect theory function with two reference points instead of one.

When one's current level is less than one's aspiration level for gains, it is hypothesized that people will naturally frame their decision in terms of gains (to reach their goals). Here, the value function is convex as is illustrated in Figure 3 as well. Hence, one should be risk seeking in those situations since increases in gains are valued more highly than decreases in gains (losses) for that part of the curve. The amount of risk seeking should depend on the distance between the current level and the aspiration level: the smaller the distance, the greater the risk seeking due to the convexity of the function.

If a person's current level is equal or greater than the aspiration level, it is hypothesized that this person will spontaneously frame his or her decision in terms of losses (to avoid falling below their desired goal state).

Here, the value function is concave. Hence, one should be risk averse since increases in gains are valued less highly than decreases in gains (losses) for that part of the curve. The amount of risk aversion should also depend on the distance between the current level and the aspiration level: the lesser the distance (and hence the closer one is to falling below the aspiration level), the greater the risk aversion due to the concavity of the function.

When faced with losses, if a person's current level is greater than the avoidance level, then it is hypothesized that this person will naturally frame his or her decision in terms of losses (to avoid reaching the avoidance level). Here, the value function is concave. Hence, one should be risk averse in those situations since the positive utility of decreasing one's losses should be less than the negative utility of increasing one's losses. The amount of risk aversion should depend on the distance between the current level and the avoidance level: the smaller the distance, the greater the risk aversion due to the concavity of the function.

If a person's current level is equal or less than the avoidance level, it is hypothesized that this person will naturally frame his or her decisions in terms of gains (to get above the avoidance level). Here, the value function is convex. Hence, one should be risk seeking since the positive subjective utility of decreasing one's losses should be greater than the negative subjective utility of increasing one's losses. The amount of risk seeking should also depend on the distance between the current level and the avoidance level: the smaller the distance, the greater the risk seeking due to the convexity of the function.

The threshold nature of the aspiration levels and avoidance levels leads to systematic effects on framing of expectations. As people move closer to their avoidance level, it is proposed that in some cases, the absolute value of this level increases so that people mentally prepare themselves for bigger losses while accommodating current losses. This hypothesis is consistent with findings by Loewenstein and Linville (1986) who find that people lower their expectations as the possibility of a loss approaches. This is also similar to the adaption level discussed by Helson as early as 1948 (Helson, H., 1948). This cognitive shifting of the avoidance level would reduce risk aversion in loss situations because the magnitude of the loss the decision maker is willing to accept is increasing. These two effects should occur in domains where outcome level is a continuous variable such as money. In other situations where outcome can be viewed as a threshold type, dichotomous variable e.g. graduating from college versus not graduating, winning a game versus losing, the aspiration and avoidance levels should not change over time. In the military situation where mission success is an aspiration level, the definition of mission success is potentially open to change in response to changing data or conditions leading to a reframing of the situation.

One could argue that prospect theory can incorporate the present framework, if one assumes that the reference point is equal to either the aspiration level for gains or the avoidance level for losses. If this is done, then the value function will be concave for points above the reference point and convex for points below it as predicted in the proposed framework. Similarly, the value function would be convex for points above the avoidance level and concave for points below it as predicted here. There are other factors which distinguish between the current framework and prospect theory, and these will be briefly discussed.

One issue is locus of control. Prospect theory generally described gambles or outcomes where the locus of control was external. The decision maker could not alter the probability of the component outcomes of the gamble. For the current framework, for outcomes whose locus of control is internal, gains below the aspiration level would be viewed by people not as gains but as reductions in losses (since if the aspiration level is the reference point, all values below it are losses). This implies that a middle manager who wishes to be company president would view a promotion to vice president as a reduction in losses rather than a gain. This does not seem plausible.

Again, the issue becomes one of aspiration level and continuity of the outcome continuum. For military decisions, the status quo could be a desirable point or not depending on the aspiration level. Suppose the aspiration level is to take objective A and hold for time period t where A is 10km or so ahead (say east). The status quo location (ignoring other details of the aspiration level) is below the aspiration level. The commander looks at options that will achieve the aspiration level. Any gains short of A, though not losses in reality are losses conditional on the mission. Thus, the commander considers all options conditional on objective A as a reference point. The current status quo would be represented as a loss at time t. Two days ago it was a gain. The commander thus must evaluate options on an "as if" basis. If I were to achieve A, which would be better, to move forward 20km to the International German Border (IGB) and lose 20 more personnel in doing so, or to stay at A and lose no one. Then he considers a second set of options. If I have the choice of being 5km west of A with a loss of only 70 men total or at A with a loss of 370, which do I prefer? In this case, he will prefer the option involving mission achievement. The tradeoffs between personnel lost and terrain achieved change depending on position with respect to mission achievement. Given that the mission is not achieved, the commander will expend (pay or cost) many personnel to move the 5km to achieve the objective. Given mission attainment, he will not expend one tenth the personnel to move four times the distance, even if such movement is consistent with mission goals (i.e., is not negative with respect to desired overall mission goals.)

The reference points with respect to goals are points of discontinuity. As earlier indicated, the commander has a "lexicographic-like" value function conditional on his position with respect to different reference points. The military situation represents one end of a continuum with respect to the severity of the described discontinuity. Practical constraints prohibit him sacrificing all personnel to achieve the mission, and the situation is more complex and dynamic than the simple one presented here. Still there are several important issues here worth reiteration.

- The commander is considering gains and losses with respect to a reference point, an aspiration level which is not his status quo. This differs from the prospect theory situation.

- The commander has some control over the probability of the outcomes and this affects his perceptions of uncertainty and the associated subjective probabilities. By limiting the scenarios considered to those involved in planning mission success, he will potentially succumb to biases that result from the availability and representativeness heuristics. These involve evaluating options with respect to the most cognitively available scenarios,

and assessing event probabilities in terms of representativeness of a class of events under consideration. If he is on the offense, he may evaluate scenarios in which the enemy is on the defensive and thus reacts to his offensive move according to the commander's understanding of enemy doctrine. He may discount or ignore scenarios in which the enemy has anticipated his actions and is instituting a novel defense or even is on the offensive himself. The commander is thus likely to overestimate his probability of success in such situations in which he is on the offensive. This would be especially true for cases where he was fairly certain about the situation. In situations of ambiguity, commanders often worst case their planning. That is, of the available things the enemy might do, the commander assumes that the enemy chooses the worst action from the commander's perspective, and so the commander plans for that. Depending on his ability to counter the enemy in this case, he could over or under estimate the probability of a favorable outcome.

There are many considerations here. The important point is that in the military situation, the commander's probabilities depend on actions he plans, and assuming independence of probabilities and values is likely to be incorrect.

- There are critical dependencies between outcomes. Trying to attain the mission can increase the probability of high losses and thus outcomes near the avoidance level. Redefining mission success can potentially reduce the probability of losses near the avoidance level.

- The commander's preferences will likely indicate that he has a discontinuous value function, for tradeoffs above and below the aspiration level will be quite different.

- Plans are made from the status quo, but outcomes are evaluated from the positive (and negative) reference points.

Hypotheses to be tested

The present study investigates portions of the above theoretical framework. Of interest is decision making in "gain" situations where potential outcomes are either above or below the aspiration level and "loss" situations where potential outcomes are either above or below the avoidance level. The results of the present study will have implications for the shape of the value function in the proposed revised prospect theory.

There are five specific hypotheses to be tested in Study 1. Hypothesis 1 concerns choices among potential gains when the current level is below the aspiration level. It is hypothesized that decision makers will prefer a gamble that has a 50-50 chance of either reaching the aspiration level or making no progress to a sure gain equivalent to half the distance between the current level and the aspiration level.

Hypothesis 2 concerns choices among potential gains when the current level is at the aspiration level. It is hypothesized that decision makers

will prefer a sure gain over a gamble that has a 50-50 chance of no gain or a gain that is twice the size of the sure gain.

Hypothesis 3 concerns choices among potential losses when the current level is above the avoidance level. It is hypothesized that decision makers will prefer a sure loss equivalent to half the distance between the current level and the avoidance level over a gamble that has a 50-50 chance of no loss or a loss that puts decision makers at the avoidance level.

Hypothesis 4 concerns choices among potential losses when the current level is at the avoidance level. It is hypothesized that decision makers will prefer a gamble that has a 50-50 chance of either placing the decision maker above the avoidance level or incurring further losses to an option with a sure outcome of remaining at the avoidance level.

Hypothesis 5 concerns the relative magnitude of the effects for gain and loss situations. Both Kahneman and Tversky's (1979) prospect theory and the current framework hypothesize that the value function is steeper for losses than it is for gains. Therefore, the shift from risk seeking to risk aversion should be greater as decisions shift from involving outcomes below and above the avoidance level versus below and above the aspiration level. In other words, the effects for hypotheses 3 and 4 are hypothesized to be greater than the effects for hypotheses 1 and 2.

METHOD

Participants

There were 55 participants in the present study. Participants ranged in rank from Chief Warrant Officer 2 (CW2) to lieutenant colonel (LTC). All participants were active duty officers from either Ft. Carson, Ft. Polk or Ft. Hood. Participants were from a variety of branches including combat arms, military intelligence (MI) and combat support.

Materials Used

Six problems were constructed in paper and pencil format. The first four problems represented abbreviated military scenarios involving a Soviet threat in Germany. In each of these problems, participants were asked to role play the commanding general. The last two problems represented choices among monetary gambles.

In problem 1, the participants were told that they were on the offensive (after initial defensive operations) and stood some sixty kilometers from the inter-German border (IGB). They were given two courses of action to chose from, one that achieved 30 kilometers for sure, and one that had a 50-50 chance of making no progress or reaching the IGB. Participants were instructed that reaching the IGB was equivalent to winning the war (which was their assumed aspiration level).

In problem 2, the participants were told that they were on the defensive and stood some sixty kilometers from the Atlantic Ocean ports. They were given two courses of action to chose from: pulling back and giving up 30 kilometers and holding thereafter for sure; and holding the current position with a 50-50 chance of successfully holding the position or allowing a penetration that would enable the enemy to seize that Atlantic ports. Participants were instructed that allowing the enemy to seize the Atlantic ports was equivalent to losing the war (which was their assumed avoidance level).

In problem 3, the participants were told that they were on the offensive (after initial defensive operations) and stood on the IGB. The enemy was suing for peace and the friendly forces were looking to strengthen their bargaining power. Participants were given two courses of action to chose from, one that achieved 30 kilometers for sure, and one that had a 50-50 chance of making no progress or seizing 60 kilometers.

In problem 4, the participants were told that they were on the defensive and were pushed back and lost the Atlantic Ocean ports. Surrender by the friendly forces was imminent. Participants were given two courses of action to chose from: surrender for sure or counterattack with a 50-50 chance of successfully regaining the ports and continuing to fight or risking defeat and new enemy retaliatory initiatives.

Problem 1 was constructed to test hypothesis 1; Problem 3 was constructed to test hypothesis 2; Problem 2 was constructed to test hypothesis 3; and Problem 4 was constructed to test hypothesis 4. The remaining two problems were constructed to replicate prospect theory's findings that decision makers

are risk averse with respect to gains and risk seeking with respect to losses for gambles with no explicit aspiration level or avoidance level.

In problem 5, participants were offered a (hypothetical) choice of a sure $500 or a 50-50 chance of $1,000 or nothing. In problem 6, participants were offered a choice of a sure loss of $500 or a 50-50 chance or no loss or a loss of $1,000.

Procedure

All six problems were administered in paper an pencil format the order described above. Participants were instructed to read through each problem sequentially and choose the option that appealed to them. Participants were instructed to treat the probability information (i.e., "sure thing" and "50-50") as reliable. Post-session discussions indicated that participants did so. Participants were allowed to work at their own pace and were assured that there were no "right" or "wrong" answers to the problems--that it was merely a survey of preferences.

RESULTS

Of the 55 participants used in the present study, one circled both options on one of the problems, therefore making the data for that problem unusable. The fact that the participant did that for one problem calls into question whether he understood the procedure. Therefore all data from that participant is excluded from the analysis, leaving 54 observations for each problem.

Table 1 presents the number of participants who selected the sure thing outcome (risk averse) and the gamble (risk seeking) for each of the six types of problems. The "monetary" problems were those designed to replicate the Kahneman and Tversky findings.

Table 1

Number of Participants Choosing Sure Thing and Gamble for Each Type of Problem

Monetary Problems	Sure thing	Gamble
Gain	36	18
Loss	10	44

Replicates Kahneman and Tversky

Military Problem	Sure thing	Gamble
Gain		
Below Aspiration Level	21	33
Above Aspiration Level	33	21
Loss		
Above Avoidance Level	39	15
Below Avoidance Level	7	47

As can be seen from Table 1, the monetary problems did indeed replicate the Kahneman and Tversky findings. When choosing among a sure $500 and a 50-50 chance at $1,000 or nothing, 36 of 54 or two-thirds of the participants chose the sure thing. Participants in this problem were risk averse as in the Kahneman and Tversky findings ($Z = 2.31$, $P = .01$).

Similarly, when choosing among a sure loss of $500 or a 50-50 chance at a loss of $1,000 or no loss, 44 of 54 or 81% chose the gamble ($Z = 4.49$, $P < .001$). Participants in this problem were risk seeking as in the Kahneman and Tversky findings. These two results suggest that although the present study involved participants who are active duty military officers (and thus perhaps not a random population sample), their results were no different than that of a more general population.

However, when outcomes are described relative to an aspiration level or an avoidance level which is more like the military situation, the results are different than the standard prospect theory findings and are consistent with the proposed extension of prospect theory.

For gains where the current level is below the aspiration level, 33 of 54 participants or 61% chose the gamble which had a 50-50 chance of reaching the aspiration level or no gain over a sure gain below the aspiration level ($Z = 1.50$, $P = .067$). This supports hypothesis 1 and suggests that participants are risk seeking for gains below the aspiration level.

On the other had, for gains above the aspiration level, 33 of 54 participants (61%) chose the sure gain over a 50-50 chance to get a larger gain or no gain ($Z = 1.50$, $P = .067$). These participants had become risk averse, thus supporting hypothesis 2.

One way to validate that participants are risk seeking for gains below the aspiration level and risk averse for gains above the aspiration level is to look at their individual data. Of the 54 participants, 10 were risk averse for both types of problems and ten were risk seeking for both types of problems. The participants of most interest however, are those who were risk seeking for gains below the aspiration level and risk averse above and those who showed the opposite tendency, namely, risk averse below and risk seeking above. Of the 34 remaining participants, 23 showed the expected pattern, while only 11 showed the opposite pattern. This difference is statistically significant using a binomial distribution that assumes and equal probability for each pattern, $p = .029$, one-tailed.

For losses above the avoidance level, 39 of 54 participants (72%) chose the sure loss above the avoidance level over a gamble that had a 50-50 chance of either reaching the avoidance level or leading to no further losses ($Z = 3.13$, $P < .001$). This supports hypothesis 3 and suggests that participants were risk averse for losses above the avoidance level.

For losses below the avoidance level, 47 of 54 participants (87%) chose the gamble that had a 50-50 chance of either moving above the avoidance level or sinking even further below it over remaining at the avoidance level for sure ($Z = 5.31$, $P < .001$). This supports hypothesis 4 and suggests that participants were risk seeking for losses at or below the avoidance level.

Again, a comparison can be made at the individual level to examine the tendency for participants to exhibit the predicted affects. Of the 54 participants, 7 were risk averse for both problems and 15 were risk seeking for both problems. Again, the participants of most interest are those who were risk averse for losses above the avoidance level and risk seeking below and those who showed the opposite tendency, namely, risk averse below and risk seeking above. Of the 32 remaining participants, all showed the expected pattern. This difference is statistically significant using a binomial distribution that assumes and equal probability for each pattern, $p < .0001$.

These results support the four major hypotheses regarding the extension of prospect theory. Specifically, these results support the contention that the value function is convex for gains between the current level and the aspiration level (thus leading to risk seeking behavior), concave for gains above the aspiration level (thus leading to risk averse behavior), concave for

losses between the current level and the avoidance level (thus leading to risk averse behavior), and convex for losses below the avoidance level. Alternately, these results support the three reference point version of prospect theory that differentiates between goal directed, risky choices, and non-goal directed choices.

The fifth hypothesis to be tested in the present study is that the shift between risk aversion and risk seeking around the avoidance level is greater than that around the aspiration level, suggesting that the value function is steeper for losses rather than gains. The general pattern of data bears this trend out. Specifically, for gains, whether the outcomes lie above or below the aspiration level produces a swing of 22% in the number of participants who prefer a sure gain over a gamble (39% for gains below the aspiration level and 61% for gains above the aspiration level). On the other hand, for losses, whether the outcomes lie above or below the avoidance level produces a swing of 59% in the number of participants who prefer a sure loss over a gamble (13% for losses below the avoidance level and 72% for losses above the avoidance level).

An additional test of this effect involves looking at the individual participants' data. In particular, one can look at the number of participants who tended to move toward versus away from the hypothesized effects for both losses and gains. Table 2 below presents a matrix of the participants who shifted in the direction of the hypothesis, away from it or stayed the same (i.e., picked the sure outcome or the gamble in both of the gain or loss problems) by type of problem gain or loss.

Table 2

Number of Participants who Moved in Direction Toward or Away from the Predicted Effect or Made a Consistent (same) Choice.

	Gain Problems		
	Opposite Effect	Predicted Effect	Same
Loss Problems			
Opposite Effect	0	0	0
Predicted Effect	6	15	11
Same	5	8	9

Of interest in Table 2 is the number of participants who either shifted in the hypothesized direction (moved from risk averse to risk seeking as outcomes go below the aspiration or avoidance level) for one type of problem (gain or loss) while remaining the same (risk averse or risk seeking) or moving opposite the hypothesized direction (i.e., moving from risk seeking to risk averse as outcomes go below the aspiration or avoidance level) for the other type of problem (loss or gain) and the number of participants who remained the same for one type of problem while moving away from the hypothesized direction for the other type of problem.

In other words, of interest is the participants who more closely matched the hypothesis for losses than for gains versus those who showed the opposite effect. More closely matching the hypothesis could be accomplished by either having the participant show the hypothesized effect for losses but not for gains or by having a consistent strategy for losses (either risk seeking or risk averse) but a strategy that was opposite to the hypothesis for gains (i.e., risk seeking above the aspiration level and risk averse below).

The opposite effect to hypothesis 5 would occur when the participants were closer to the effects predicted by hypotheses 1 and 2 than they were to 3 and 4. This would happen when the participant either showed the hypothesized effect for gains but not for losses or by having a consistent strategy for gains (either risk seeking or risk averse) but a strategy that was opposite to the hypothesis for losses (i.e., risk seeking for losses above the avoidance level and risk averse below).

We look at each of these components (i.e., participants whose choices tended to be consistent with hypothesis 5 versus those who were the opposite) by examining the data in Table 2. Table 2 shows that there were six participants who showed the hypothesized effects for losses but showed the opposite effects for gains, eleven who showed the hypothesized effects for losses but were consistent in their gain strategy (either always risk averse or risk seeking), and five who used a consistent strategy for losses but showed opposite the hypothesized effects for gains. The total of these three cells is 22--the number of individual participants whose choices were in the direction predicted by hypothesis 5.

For those who went against the hypothesis, there were eight participants who showed the predicted effects for gains, but used a consistent strategy for losses. There were no participants who showed the opposite effects that were hypothesized for losses (risk seeking above the avoidance level and risk averse below).

Therefore, in the individual data, there were 22 participants whose results tended to go in the direction of hypothesis 5, while there were eight whose results went in the opposite direction. This difference is statistically significant using a binomial test, $p = .008$, one-tailed. Thus, the individual data provided additional support for hypothesis 5.

DISCUSSION

The data in Study 1 provide overwhelming support of the proposed extension of prospect theory to goal directed decision making. Specifically, the results suggest that the value function is convex for gains between the current level and the aspiration level (thus leading to risk seeking behavior) and concave for gains above the aspiration level. The results also suggest that the value function is concave for gains between the current level and the avoidance level and convex for gains below the avoidance level. The present results also confirmed prospect theory's contention that the value function is steeper for losses than for gains.

As discussed in the introduction section, the present reformulation of prospect theory's value function to one involving three reference points can be reconciled with the original prospect theory value function by taking into account that the latter did not take specific goals into account. If there are no aspiration and avoidance levels, then the revised value function would collapse into the original value function.

The revised value function can be a useful tool in understanding the types of courses of action and levels of risk military commanders will take in order to seize objectives or accomplish their mission. The revised value function also underscores the importance, when attempting to predict his behavior, of understanding what the commander's goals are, both with regard to gains and losses, given that decision making changes dramatically around the aspiration and avoidance levels. If this information is inaccurate, predictions could be greatly in error.

This argument applies to understanding the enemy's goals (his aspiration and avoidance levels) as well. A misunderstanding of where the enemy's aspiration and avoidance levels could prove particularly costly when trying to infer what the enemy is likely to do. As a result, the importance of the intelligence collection and analysis process is emphasized as this is the primary mechanism for inferring enemy goals and intentions. Normally, intelligence efforts are focused on understanding the enemy's most probable course of action. However, the present theoretical framework suggests that the underlying aspiration and avoidance levels are crucial in such inferences, since knowledge of them will help in the prediction of how much risk the enemy is willing to assume to reach his specific objectives. Knowing when the enemy will assume great risk and when he will be cautious is obviously valuable tactical information.

The present framework also serves as a basis for communicating risk information. The commander in his guidance to the staff specifies the objectives of the course of action the staff is to develop. The staff is then left to develop a course of action, which by definition includes how much risk to assume to achieve those objectives. The present framework suggests that the commander needs a means to communicate to his staff exactly what his goals (aspiration and avoidance levels) are and how much risk he is willing to assume under different tactical scenarios. Similarly, the staff needs a means to provide risk-relevant feedback regarding potential courses of action to help the commander assess what levels of risk he needs to assume in order to have a given likelihood of achieving his goals.

The present framework would also play a valuable role in the development of tactical decision aids. A good decision aid not only helps the user see the likely outcomes of different options being considered, but also helps that user understand how well those options will fare in meeting the user's goals. It is argued that capturing the value of those goals can be done using the concepts underlying the revised prospect Theory value function. For example, in a wargaming/course of action analysis aid, the revised value function could serve as a basis for assigning value to outcomes of different courses of action. The aid could then present the user with the risks and resources required to achieve these different levels of value.

The present framework holds promise to better the understanding of how decision makers choose courses of action and also holds promise in aiding those decision makers, but the present study is far too theoretical to be immediately applicable to the tactical decision making context. The next study provides a more realistic demonstration of how the theoretical framework could be used to derive specific hypotheses regarding tactical decision making under uncertainty and time stress that can be used to predict decision making behavior under those conditions.

STUDY 2

INTRODUCTION

Study 1 provided support for the present formulation of prospect theory. In particular, Study 1 showed that people are risk seeking for gains below the aspiration level, but risk averse for gains above the aspiration level and risk averse for losses above the avoidance level, but risk seeking for losses below the aspiration level.

However, Study 1 was conducted using rather abstract military problems that do not capture the richness of the problems that real-life decision makers might encounter, as well as the factors of time stress and uncertainty that are inherent in the normal decision process. The purpose of Study 2 is to test the present theoretical framework further under more realistic scenario conditions and to look at how time stress and uncertainty impact the decision process. Of particular interest are impacts with respect to the important military situation assessment and course of action development.

The remainder of this section develops new general hypotheses derivable from the extension of prospect theory. These in turn are used to select four specific hypotheses that are tested in Study 2.

Outcome Distributions.

The traditional decision analytic decision tree representation of acts and its subsequent act-event combinations eventually links any decision to a set of outcomes with associated probabilities. In fact, the decision tree serves as a mechanism for formalizing the relationship between the decision action and the distribution of subsequent outcomes. By collapsing a decision tree, an action can be characterized by a distribution of outcomes, with associated probabilities. Each outcome has attached to it the sum of the (normalized) probabilities of the decision tree branches for which it is the resultant endpoint outcome. This is a convenient theoretical framework to characterize an ideal planning representation of combat decisions, for it provides a mechanism that yields a distribution of outcomes with associated, well-defined probabilities. It also provides a concept for a commander's cognitive model with respect to available actions and outcomes. Using this formal theory as a basis for the cognitive model, it is quite clear that the commander will not have the formal decision tree as a well-defined mental model of his planning situation. Constructing decision trees by analyzing decisions, and in doing so, extracting acts, events, outcomes, probabilities, and utilities is a complex, tedious process for even fairly simple decisions. For complex ones, the tree is clearly described by the often used phrase "a bushy mess", and the definitive connection of acts and events with subsequent outcomes is often tenuous due to the need for simplification in representing potential scenarios. Thus, two things are likely to be problematic. One is the definition of the outcome of a specific decision tree branch. Second, and a logical consequence of the first, is the problem that probabilities of specific outcomes will be ill-defined even for outcomes that can be specified as endpoints for tree branches. This is due both to difficulties in identifying all branches for which a specific outcome might be a consequence and to

assessing aggregated probabilities of the associated sequences of events for a specific branch.

Thus, although the commander could be used for the judgments required for a skilled decision analyst to develop a decision tree, he will not have a clearly defined decision tree as a mental model. Act-event-consequence combinations will be loosely defined and outcomes will have probabilities of varying degrees of precision associated with them.

Recalling the goal directed nature of military planning discussed in Study 1, the cognitive representation of uncertainty in outcomes can be thought of as a mental representation of the outcome continuum of Figure 3 of Study 1 with a probability distribution defined over outcomes denoted as outcome distributions for specific options available to the commander. Assuming the utility function depicted in Figure 3 as being representative of the goal-directed nature of the commander's thinking, different cases can be postulated.

Several interesting cases are illustrated in Figure 4 which displays the prospect Theory Utility function and the locations of several options with respect to it. Consider options $O_1 = (a, \frac{1}{2}, d)$ and $O_2 = (b, \frac{1}{2}, c)$. these are both symmetric around their common expected value of $e = \frac{1}{2}(a + d) = \frac{1}{2}(b + c)$. These options both have outcomes above the Avoidance Level but below the reference point status quo where the curve is concave. In this case the certain option e has the highest utility ($h = U(e)$) of options e, O_1, and O_2. The expected utilities of O_1 and O_2 are located one half the distance along the line segment joining the utilities of the components of the options. Thus the utility of the expected value or certain option exceeds the expected utility of options having the same expected value, meaning risk aversion. Figure 4 shows a similar set of relations for options O_3 and O_4 having outcomes above the aspiration level and having expected value n. Above the aspiration level and above the avoidance level (but below the status quo as reference point), the decision maker having the illustrated utility function will be risk averse.

In Figure 5, the options O_5 and O_6 have the same expected value v, but O_5 is riskier than O_6. Note that it also has an outcome e.c.y that is right at the aspiration level. In this case, O_5 is preferred to O_6 which is preferred to the certain option v which equals the common expected value of O_5 and O_6. In this region of convexity, the decision maker will be risk seeking. The same will be true for the region below the avoidance level.

As indicated in Study 1, the curve shows the preferences of an individual who chose options on the basis of expected utility and who had the revised prospect theory value function. It has already been noted that representing the outcome set as a single continuum is a simplification. Even if it were a vector which is an aggregated multi-attributed utility function, it is somewhat unlikely that it would be clear and/or always used consistently by decision makers. Also the steepness of the function is uncertain. As has been discussed, the aspiration and avoidance levels may serve as thresholds, and behavior, with respect to them will be conditional and potentially discontinuous. For example, the utility for levels of casualties will vary

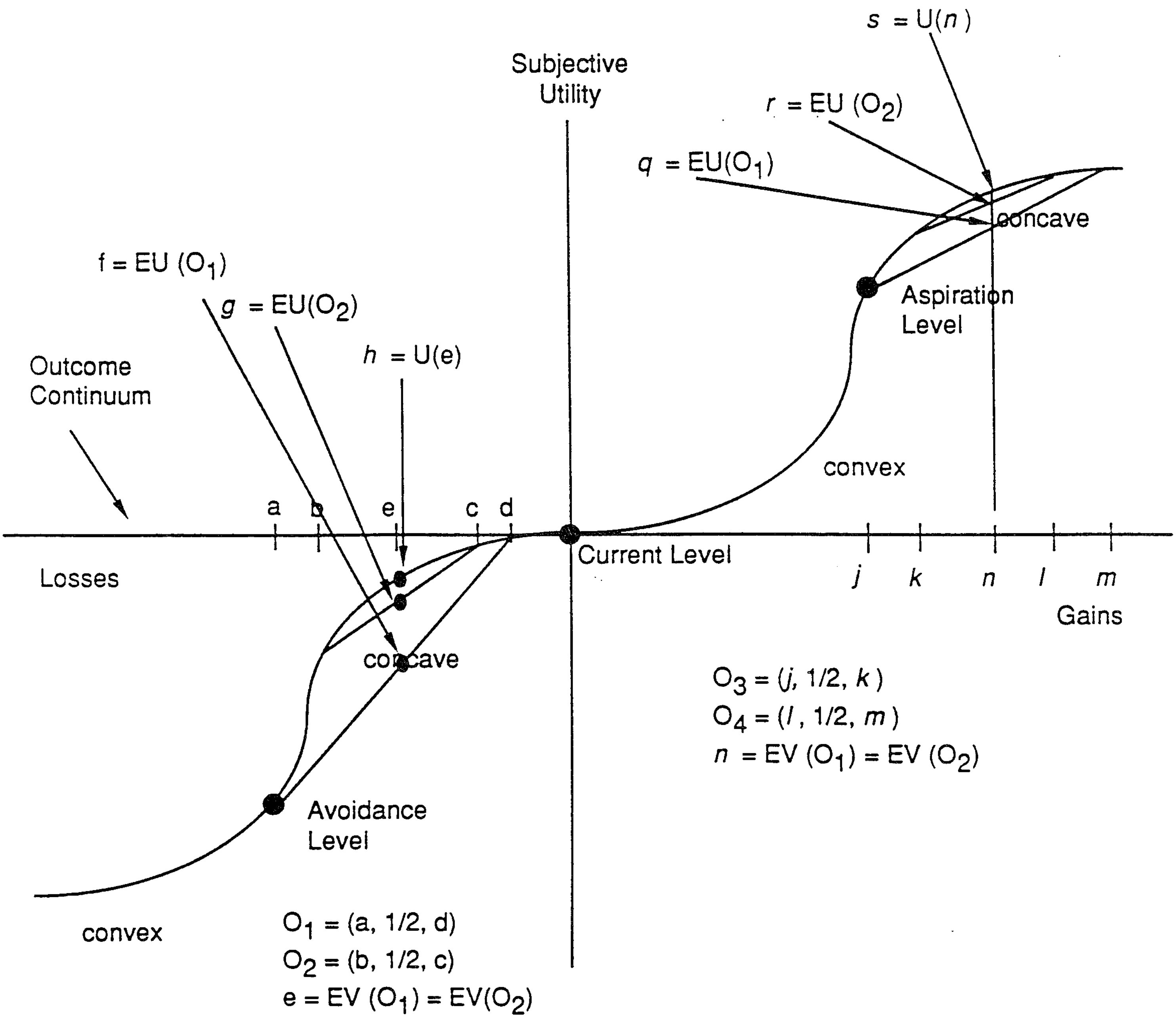


Figure 4. Using the prospect theory utility function to display choices among several options.

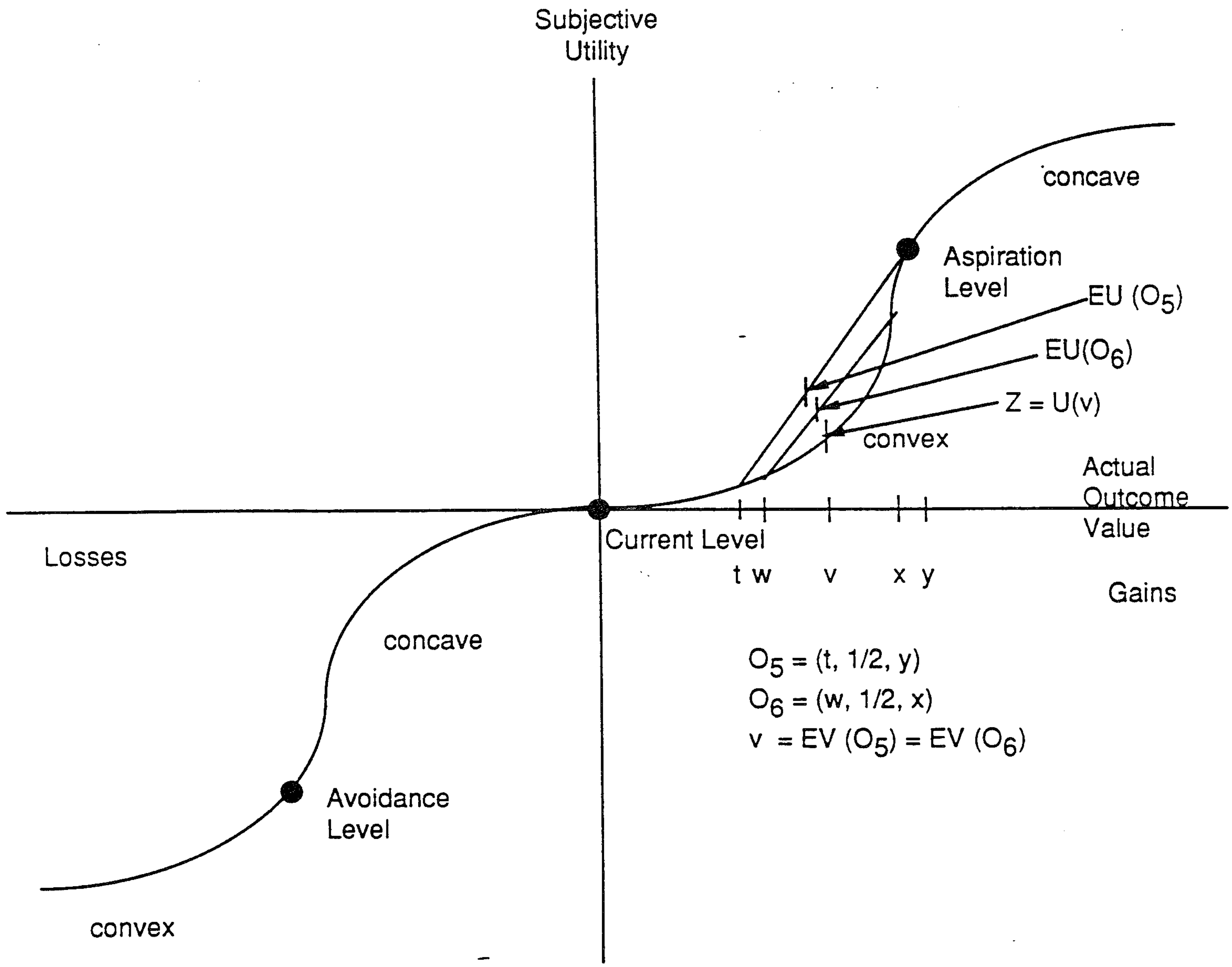


Figure 5. Using the prospect theory utility function to display choices among options below the aspiration level.

dramatically by whether the utility is conditional on achievement of aspiration level, on being above but near avoidance level, etc.

Continuing the discussion, consider the case when outcome distributions straddle the aspiration level, i.e., some outcomes lie above and some lie below. Then the hypothesized behavior becomes more difficult to predict. A significant factor is the relative degree of convexity of the value function below the aspiration level versus the degree of concavity of the value function above the aspiration level. For example, if the two functions were symmetrical, one would expect indifference between a set of options whose outcome distributions have midpoints at the aspiration level but have different variances. However, it is hypothesized that for most people, the value function is steeper for the convex portion of the curve than it is for the concave portion (this is similar to prospect theory's contention). Hence, stretching an outcome distribution above and below the aspiration level should have a resulting negative effect on its subjective utility. As a result, decision makers are uncertainty/ambiguity averse when outcome distributions are symmetrical around the aspiration point. Shifting this distribution higher (so that the midpoint moves above the aspiration level) will increase the aversion to uncertainty (i.e., wide range of outcomes), while shifting this distribution lower (so that the midpoint is below the aspiration level) will decrease the aversion to uncertainty and at some point the aversion should become seeking (e.g., as when the distribution lies entirely below the aspiration level--see above). (Again, these statements are made relative to the notion of a single outcome continuum.)

Summarizing for options whose outcomes lie below the avoidance levels, it is expected that the convexity of the value function in that area will cause decision makers to be ambiguity/uncertainty seeking, much the way they are hypothesized to be on the convex portion of the value function below the aspiration level.

For options whose outcomes lie between the avoidance level and the status quo reference point, it is expected that the concavity of the value function in that area will cause decision makers to be ambiguity/uncertainty averse, much the way they are hypothesized to be on the concave portion of the value function above the aspiration level.

For options whose outcomes straddle the avoidance level, again the attitude that decision makers have toward the range of the outcome distribution is intimately linked to the relative convexity of the value function below the avoidance level versus the concavity of the value function between the reference point and the avoidance level. As discussed, it is hypothesized that the curve is steeper above the avoidance level (i.e., toward the reference point) than it is below. Hence, decision makers are uncertainty/ambiguity seeking when outcome distributions are symmetrical around the avoidance level. Shifting this distribution higher (so that the midpoint moves above the avoidance level toward the reference point) will decrease the preference for uncertainty (i.e., wide range of outcomes), while shifting this distribution lower (so that the midpoint is below the avoidance level) will increase the preference for uncertainty. As the distribution moves higher, at some point the preference for uncertainty should become aversion (e.g., as when the distribution lies entirely above the avoidance level--see above).

Another interesting case is where the outcome distribution lies across the status quo reference point. Here, the preference/aversion to uncertainty would depend on the relative steepness of the convex portion of the gains curve versus the concave portion of the loss curve. Prospect theory argues that losses loom larger than gains, suggesting that the value function is steeper for losses than it is for gains. There seems to be no reason to argue differently, hence it is hypothesized that decision makers are uncertainty/ambiguity averse for outcome distributions that straddle the reference point. Here, the amount of aversion increases as outcomes shift toward the avoidance level and decreases as they shift toward the aspiration level until the uncertainty/ambiguity aversion becomes uncertainty/ambiguity seeking.

Option Generation and Situation Assessment. Military decision making deals with uncertain outcomes where the enemy often tries to increase uncertainty through use of deception, etc. Also the outcomes associated with option typically involve high stakes. High uncertainty with high stakes defines a risky decision, and the commander has two ways to reduce risk. He can reduce uncertainty through information management, and he can reduce the consequences of facing that uncertainty by designing and evaluating the options he has for action, defined here as option generation or modification. Information management involves data collection, reduction, assimilation, and interpretation, all required to reach an assessment of the situation denoted as situation assessment (SA). By increasing the accuracy of the situation assessment, the uncertainty of the action/event combination is reduced. By conducting a comprehensive option design and evaluation, the commander better controls the outcome associated with the action/event combination. By increasing the accuracy of the situation assessment, he provides for better option generation. He also reduces risk associated with uncertainty. Note that he may actually confirm a very bad situation, thus confirming potentially devastating act/event combination. However, uncertainty is reduced, and the potential for conditional optimization through option generation or revision is increased.

Earlier, hypotheses were discussed regarding what options decision makers will select given the outcomes that may be associated with them. Combat decision making is dynamic, and there will be times when decision makers will actively try to generate new options (or modify existing options). In particular, it is hypothesized that option generation will primarily occur when outcomes distributions lie below the aspiration level in situations conditionally focused on gains and when outcome distributions lie below the avoidance level in situations where decision makers are conditionally focused on losses. In both cases, it is hypothesized that decision makers will be trying to shift the outcome distributions so that they lie above the aspiration and avoidance levels, respectively.

It is hypothesized that there are two means by which decision makers can shift expected outcome distributions. First, decision makers can attempt to resolve uncertainty (e.g., perform situation assessment), thereby narrowing the expected range of possible outcomes. This strategy would be most useful when outcome distributions are asymmetric around aspiration and avoidance levels with the majority of the distributions lying above those points. Second, the decision makers can engage in option generation/modification. We argue that there are two such modes of option generation: risk reduction and opportunity increment. Risk reduction refers to eliminating possible outcomes which fall below aspiration and avoidance levels. Opportunity increment

refers to adding possible outcomes which lie above aspiration and avoidance levels.

Hypothesizing the value function of Prospect theory and knowing the general range of expected outcomes will enable the generation of hypotheses as to when people will modify options using opportunity incrementing or risk reducing strategies (in the section below on time stress, hypotheses regarding when decision makers will resort to situation assessment versus option generation will be discussed). When outcomes straddle the aspiration point but the majority of the outcomes are above the aspiration point, then adding higher value outcomes has little increased value due to the concavity of the curve above the aspiration level. However, since the value function is particularly steep below the aspiration level, removing outcomes which fall below the aspiration level carries a high subjective utility. Hence, it is hypothesized that in these cases, decision makers will engage in risk reducing or conservative option generation.

When outcomes straddle the aspiration level but fall largely or entirely below it, then option generation should shift to opportunity increment. In such cases, the outcomes with the lowest value are at the shallowest portion of the convex part of the value function for gains. Hence, there is little subjective utility for eliminating those outcomes. However, outcomes near the aspiration level are on a steeper part of the curve (even if they are just over the aspiration level and the concave portion of the curve hasn't leveled out yet). Hence, there is greater subjective value in generating options whose outcomes have higher value outcomes in its distribution. It is therefore expected that under these conditions, decision makers would be willing to generate riskier options.

However, once these options with higher value outcomes are generated (and the outcome distribution is less skewed across the aspiration level), decision makers may switch to a risk reducing strategy such as situation assessment or contingency planning in order eliminate options with outcomes below the aspiration level.

For outcome distributions which are symmetric around the aspiration level, the tendency to engage in risk reducing option generation versus opportunity incrementing option generation depends on the relative steepness of the value curve on its concave and convex portions. Since it was previously argued that the value curve is steeper on the convex portion, greater value is gained by eliminating the outcomes below the curve compared to adding outcomes above the curve. Hence, decision makers are hypothesized to engage in risk reducing option generation or situation assessment.

With respect to outcomes near the avoidance level, recall that the value curve is convex below the avoidance level and concave above. Hence, when outcome distributions are symmetric around the avoidance level or lie largely above it, decision makers are hypothesized to engage in risk reducing option generation (eliminating outcomes below the avoidance level). However, when outcome distributions lie mostly or entirely below the avoidance level, decision makers are hypothesized to engage in opportunity increment option generation (since the value function is steeper near the avoidance level than it is on the tail of the convex portion of the value function below the avoidance level). Again, once the decision maker has generated options with outcomes

above the avoidance level, he may switch to risk reduction option generation to eliminate outcomes below the avoidance level.

The above analysis on option selection and option generation leads to the first experimental hypothesis to be tested in the present study.

Hypothesis 1: If participants think the best course of action (COA) meets their aspiration level of success, they will focus on resolving uncertainties to verify its success. In contrast, if they think that even the best COA has little chance of success, they will generate a new course of action. In the case where the decision maker thinks that the best COA meets their projected level of aspiration, this means that the expected outcome is above the aspiration level. Further, most of all if the probability distribution over outcomes lies above the aspiration level. Conversely, when the best COA has little chance of success, this means that the decision maker's probability distribution over outcomes for this best COA lies almost all or completely between the aspiration level.

Introducing Time Stress. In a time stressed military situation, time required to design, evaluate, and implement actions is running critically short, and a crucial event that will results in outcomes is impending. Time stress requires a reduction in the time taken to process information or an increased processing rate. Similarly, time stress can involve the requirement for rapid option generation. The prospect theory extension can be considered with respect to time stress, and several concepts can be proposed.

Amount of time available directly influences the amount of information decision makers process and the amount of analysis they are able to perform. This influences the resultant amount of uncertainty/ambiguity still remaining in the situation, which can be characterized by the more diffuse subjective distributions of expected possible outcomes. This influences option selection and option generation behavior (along the lines of the hypotheses presented above).

According to this viewpoint, time stress influences behavior through the information search and analysis that decision makers perform. Similarly, because option generation includes at least subjective evaluation, the number of options considered and associated analytic evaluation that can be done are reduced. It is hypothesized that under time stress, decision makers will look at less information and perform less analysis (as was the case in the Leddo, Chinnis, Cohen, and Marvin, 1987, phase I study) and generate and evaluate fewer options. Hence, the amount of uncertainty regarding potential outcomes after analysis will be greater than would otherwise be the case. Wider resultant expected outcome distributions are hypothesized.

There is another potential hypothesis as to how time stress might influence perceived outcome distributions. Loewenstein and Linville (1986) conducted a study demonstrating that as subjects approached a deadline, they became pessimistic regarding an outcome that would effect them. The argument is that subjects were cognitively protecting themselves from potential disappointment by lowering their expectations. When decision makers are under time stress, they are, in effect, operating with a closer deadline than decision makers operating under no time stress. Hence, if the Loewenstein and Linville study generalizes, then one would expect time stress decision makers to be more pessimistic (due to their impending deadline) than no stress decision

makers. If this is so, one would expect a downward shift in the expected outcome distributions. This would lead to a larger percentage of time spent in option generation behavior in general. An informal review of the Phase I study session tapes suggests that time stress participants did engage in relatively more option generation than no stress participants.

Earlier, it was argued that decision makers could engage in situation assessment (uncertainty reduction) or option generation in order to affect their perceived outcome distributions. It is hypothesized that situation assessment is more likely to be emphasized under no stress conditions (compared to time stress conditions), while option generation is more likely to be emphasized under time stress conditions (compared to no stress conditions). Situation assessment is a time consuming process which involves sifting through a great deal of information and conducting thorough analyses. Hence, time stress decision makers are not likely to have the time to spend conducting situation assessment and will go directly to option generation, which is required effort. Thus, time stressed participants should spend <u>relatively</u> more effort on option generation than do non-stressed participants. We would expect non-stressed decision makers to spend relatively more time doing situation assessment.

Another argument for the results is the following. Given an uncertain situation as a starting point, it is obviously necessary to reduce uncertainty in order to design and evaluate options. If there is insufficient time for both efforts, the actions must sill be designed. Thus, the relative effort on situation assessment must be reduced, and this can be accomplished by worst casing the situation conditional on what is known to date. That is, of the potential situations currently feasible, design against the worst of those known to be likely, i.e., design to cover as many likely contingencies as possible.

In reality the situation assessment function (done by the intelligence or "INTELL" staff) and the option generation function (done by the Operations staff) proceed in parallel while interacting. However, there is a point beyond which the planned actions must be prepared for implementation and further INTELL input, unless absolutely devastating in implications, must be ignored. Thus the situations to be studied here, though somewhat artificial, are not unreasonable.

Results of Leddo et al. (1987) suggest that no stress participants did spend far more time in situation assessment than time stress participants and conducted this activity first, while time stress participants spent a lot of their time engaged in option generation. An informal revisiting of these data also suggest that time stress participants made more major modifications/revisions to the existing courses of action than did no stress participants. These concepts lead to the second hypothesis to be tested in the present study.

Hypothesis 2: Time stress participants will focus relatively less time on situation assessment and relatively more time on option generation while no stress participants will do the opposite.

There are two potential theoretical justifications for this hypothesis. First, the theory presented in this paper suggests that participants may react to time stress by having more uncertainty than the no stress participants

(since they have insufficient time to resolve it). This could result in an expectation of a wider range of outcomes to a given course of action. As the range of potential outcomes increases, so does the possibility that the outcome may fall below the aspiration level. When this happens, participants are hypothesized to be more motivated to develop new options. Second, Loewenstein and Linville (1986) find that as people approach a deadline, they become more pessimistic in their expectations. Here, time stress participants have an impending deadline throughout the study, suggesting that they may become pessimistic in the expected outcomes in their old courses of action. If this happened, then time stress participants would be more likely to generate a new course of action for of these two explanations would mediate the effect, but in the present study both predict the same effect.

If in fact hypothesis 2 is supported, then time stress participants can often form an incorrect perception of the situation. This is particularly true in cases where the situation changes over time. If time stress participants do not pick up on these changes, they are likely to stick with their initial impressions of the situation, which in turn would influence their course of action selections.

As indicated, an alternative coping strategy when unable to resolve uncertainty would be for participants to adopt a "worst case" perspective on the situation and plan accordingly. In fact, Army officers often state that this is what they are taught to do. If this coping strategy is used, we would expect that the behavior of participants who have not resolved uncertainty to be identical to that of participants who have resolved uncertainty and found that the "worst case" was confirmed.

Hypotheses 1 and 2 provide the rationale for hypotheses 3 and 4 to be tested in the present study.

Hypothesis 3: When the situation initially looks good but then subsequent information suggests that the situation is bad, then time stress participants will be more likely than no stress participants to pick an existing course of action that would have exceeded the aspiration level in the initially described "good" situation.

The rationale behind this hypothesis is that time stress participants will not have sufficient time to evaluate the situation fully and will therefore be less likely than no stress participants to discover that the situation has turned bad. Therefore, time stress participants should be more likely to pick an existing course of action that works in the perceived good situation.

Hypothesis 4: When the situation initially looks bad but then subsequent information suggests that the situation is good, then no stress participants will be more likely than time stress participants to pick an existing course of action that does exceed the aspiration level in the revised "good" situation.

The rationale behind this hypothesis is that time stress participants will not have sufficient time to evaluate the situation fully and will therefore be less likely than no stress

METHOD

Overview

The purpose of the Study 2 was to test the four hypotheses generated under the present theoretical framework in the context of tactical decision making by experienced Army officers under conditions of time stress versus no time stress and a perceived good versus bad tactical situation. Accordingly, a 2 x 2 x 8 design was employed with time (45 minutes versus unlimited), situation (initial good changing to bad versus initial bad changing to good), and participants as variables.

Each participant was assigned to one of the four experimental conditions, given a tactical scenario to work on and was asked to recommend a course of action. Of interest to the present study were the course of action recommendations made by the participants, the conclusions about the situation they arrived at, and the allocation of time during problem solving.

Participants

There were 32 participants in the present study. Participants ranged in rank from captain to lieutenant colonel. Participants comprised active duty officers from FORSCOM installations, the Army Research Institute (ARI), the Pentagon, instructors from the United States Military Academy (USMA) at West Point, and reserve officers from a Washington, D.C. reserve unit. Participants were from a variety of branches including combat arms, combat support, and military intelligence (MI).

Scenario Used

The scenario used in the present study was a modified version of the U.S. Army Training and Doctrine Command (TRADOC) common teaching scenario, obtained from ARI-Ft. Leavenworth and modified for the needs of the current study. A brief description of the scenario is presented below.

The modified scenario depicted the 16th Mechanized Infantry Division (16 MID), the participants' unit in the study, as currently being surrounded by enemy units. The 16 MID had been engaging in a supporting attack as part of the 10th U.S. Corps with a mission of restoring the Inter-German Boundary (IGB). The 16 MID was attacking along two axes with the 1st Brigade (BDE) as the main attack in the south and the 3rd BDE as the supporting attack in the north. The 2nd BDE was the division reserve. This attack had been stalled by elements of the 6th Combined Arms Army (CAA) and the 10th CAA.

The 21st Motorized Rifle Division (MRD) of the 6th Combined Arms Army (CAA) was located to the north of the 16 MID with elements in the rear of the 3rd BDE, the 15 MRD of the 10th CAA was located to the south of the 16 MID with elements in the rear of the 1st BDE, the 33rd Guards Tank Division (GTD) of the 10th CAA was counterattacking both the front and southern flank of the 1st BDE.

While this was going on, there had been reports that a full-strength tank army (TA), the 14th TA, might be moving into the 10th Corps sector. The

14th TA was comprised of three divisions and represented a considerable threat.

Other relevant features of the scenario were that the 16th MID had been given the 313th BDE, which had previously been a Corps asset, to use in their operations. The approximate strengths of the units were: the 1st and 3rd BDEs were at about 80%, the 2nd BDE at about 90%, the 313 BDE about 95%, the 21st and 15th MRDs were largely combat ineffective, the 33rd GTD was about 70%, and the 14th TA was about 95%.

Included in the scenario were a variety of information sources. First, there were both Corps (1:250,000) and Division (1:50,000) maps of the friendly and enemy situations. The maps contained overlays with friendly and enemy unit dispositions. These overlays were current as of 1:00 a.m. on September 6 (scenario time) which was five hours prior to the scenario time the participant started at (6:00 a.m. September 6 or 060600 Sep, military terminology).

There were three books that contained background information. The first book, "the Intelligence Workbook," contained the Corps fragmentary (frag) order and Division commander's guidance as of the start of the initial operations, a G-3 situation update that presented a brief history of the battle up to 060600 Sep and two tentative courses of action developed by the G-3 plans, a summary of enemy activity as up until 060100 Sep, a section on weather and terrain, and a section on enemy capabilities and vulnerabilities. The two tentative courses of action presented by the G-3 plans were to withdraw the division to the Phase line at the start of the operations and to join the supporting attack with the main attack and proceed onto the objectives.

The second book, "Staff Estimates Workbook," contained estimates by the G-1 (Personnel), G-2 (Intelligence), and G-4 (Logistics), as well as a friendly forces task organization, and a summary of friendly operations from the start of the battle. The Staff Estimates Workbook contained information that the participants could use to compute friendly force strengths.

The third book, "the Order of Battle Workbook," contained information that participants could use to estimate the strength and degree of threat of the enemy units. Contained in this book were information on enemy composition, disposition, strength (by personnel and equipment), tactics, and miscellaneous information such as training and leadership. All information in the Order of Battle Workbook was current as of 060100 Sep (i.e., was five hours old).

The information contained on the maps and in the workbooks represented information to get participants up to speed on the scenario. This information represented what participants would have accumulated had they been in an exercise from its inception rather than being inserted two days into an operation. Separate pure in addition to this background information, there was a series of messages that reported events that had transpired from 060100 Sep to 060600 Sep. These messages pertained to recent enemy activity and, when processed, depicted the current enemy situation. The messages were presented in intelligence summary (intsum) format and came from the division's intelligence and maneuver assets. These messages provided the ability to assess a situation change or confirm an initial estimate. The subject, in analyzing the messages could identify troop movements including units and

their numbers. By consulting the enemy OB, the subject could locate elements of the 14th tank army from the messages. He then had to infer that the presence of these unit elements in the sector or that they had gone south (or had not arrived) respectively confirmed or denied the situation as good. If the subject did not do this, he should not have been able to confidently locate the 14th tank army subsequent to 060100 Sep.

Both the messages and the G-3 situation update were used to implement the situation conditions. Two situation conditions were employed. The first was where participants were told (through the G-3 situation update) that the 14th TA was expected to enter the 16 MID's sector. In this condition the messages later confirmed that the 14th TA had, in fact, gone south to oppose the Corps main attack and would not affect the 16 MID's operation. This condition was labeled the "Bad to Good" condition since if the 14th TA did enter the 16 MID's sector (as the initial indications suggested), then the 16 MID would not have enough forces to carry out their mission. On the other hand, once it was learned that the 14th TA actually went south, the 16 MID would have enough forces to carry out its mission. Hence, the 16 MID's tactical situation initially looked bad, but changed to looking good.

The second situation condition was the reverse of the first. Here, the G-3 situation update suggested that the 14th TA would not come into the 16 MID's sector (in which case the 16 MID should be able to complete its mission). However, the messages later indicated that the 14 TA did come into the 16 MID's sector. Accordingly, this condition was labeled the "Good to Bad" condition.

The G-3 situation update also gave estimates of the likelihood of success of each course of action and expected casualties. This was done to help participants assess whether the expected outcomes of the initial courses of action would fall above or below their aspiration level. In order to estimate where participants' aspiration levels would be for success likelihoods of courses of action, the present materials were pilot tested. In the pilot test, participants were explicitly asked whether they had target probabilities of success they tried to achieve when developing courses of action. Results from the pilot test suggested that participants typically sought and picked courses of action that carried, in their perception, at least 70% - 80% probability of success (this level was confirmed with Study 2 participants who were also asked what their targets for probability of success were). Therefore, 80% was taken as a probability of success to meet or exceed an aspiration level for success.

For course of action 1 (withdraw), participants were told that there was an 80% - 90% chance of successfully pulling back to the original phase line, with additional casualties being in the 5% - 10% range. This estimate was given in regardless of whether the 14th TA was expected to be in the sector (since the 16th MID would be retreating from them). For course of action 2 (continue attack along main axis with 1st and 3rd BDEs), participants were told that if the 14th TA was not present in the 16th MID sector, the probability of successfully reaching the objectives was 80% with casualties in the 25% - 30% range. Participants were told that if the 14th TA was in the 16th MID sector, then the probability of successfully reaching the objectives was only 10% - 20%, with casualties expected to be as high as 50%. Post-experimental debriefing suggests that participants tended to believe these initial estimates.

Among the several purposes of these estimates was to establish that the 14th TA was key to whether participants could achieve their aspiration levels. When the 14th TA is not present, then the probability of successfully attacking is at or above participants' aspiration levels. When the 14th TA is present, then the probability of successfully attacking is far below the aspiration level.

Procedure

All participants were tested individually. The experimenter team consisted of a psychologist who collected the data (described below) and a military consultant who briefed participants on the scenario and answered questions of a military nature that participants had. The role of the military consultant was to clarify ambiguities (if any) in how participants read scenario information and to provide additional background information, when requested by participants, not presented in the scenario.

All participants received a general description of the study. They were then asked to fill out background questionnaires that contained questions regarding their experience and training. Participants were then given a background briefing by the military consultant. This background briefing consisted of the military consultant reading the G-3 Situation Update Contained in the Intelligence Workbook. Since the G-3 Situation Update contained the initial intelligence estimate of whether the 14th TA would enter the 16 MID sector, the experimental condition of situation (Good to Bad/Bad to Good) was introduced at this point. Participants were assigned randomly to one of these conditions with the constraint that half of the participants run would be in each condition.

After the G-3 Situation Update was presented, the military consultant showed the participant what information he had available to work with. This included pointing out that the map and books were five hours old and that any new information was contained in the message file.

After the materials were reviewed with the participant, the time condition was introduced. Participants were told that they would be expected to recommend a course of action, in the form of a verbal briefing, either in 45 minutes (time stress condition) or whenever the participant was ready (no stress condition). Participants were told that they could either choose one of the two courses of action presented by the G-3 plans in the G-3 Situation Update, modify them or develop an entirely new course of action. Participants were instructed to think aloud as much as they felt comfortable doing in order to articulate their decision making process and information use.

In order to facilitate this, there were two scheduled breaks in each session. The first came after participants finished reviewing the books or the messages (whichever came first) and the second came after participants finished the other. During these breaks, participants were asked to review what they had been doing and the reasoning process they were going through. Participants were also asked whether their opinions of any of the initial courses of action offered by the G-3 plans had changed. In order to prevent time stress participants from rushing through these breaks, they were told that the 45 minute clock would stop so that they would not be penalized for providing feedback to the experimenters.

Throughout the session, the psychologist experimenter monitored the participants' behaviors and collected measures on them. (In addition, the sessions were videotaped in case they were needed to verify the experimenter's data collection.) Specifically, the general categories of background information, situation assessment, course of action analysis, and option generation were defined in terms of specific observable behaviors such as reading the commander's guidance, plotting messages on the map, and wargaming courses of action. The experimenter marked when these behaviors were occurring and how much time participants spent on each of these behaviors. Later, the specific behaviors were aggregated into the four more general categories in order to evaluate the experimental hypotheses. Information use was relatively easy to assess--the experimenter noted when the participants were using the books or going through the messages. In order to assess when participants were analyzing courses of action or generating new options, the experimenter relied on the participants' think-aloud. The experimenter periodically asked the participants what they were doing, particularly when their behavior was ambiguous. All recording of time was done using a stopwatch.

After the session, participants gave a briefing on their courses of action. The experimenter then asked a series of questions including ones to determine whether participants had found the 14th TA and what types of aiding they would have wanted to help their problem solving. Other questions asked, but not included in the data analysis include the reasons behind the course of action recommendations, sources of uncertainty considered by the participants, their estimated probability of success of their recommended courses of action, and what information they would have wanted that they did not receive. Participants were then debriefed on the purpose of the study.

Data Used

There were three types of data that are relevant to the present study. The first is how participants allocated their time across information and tasks. In particular, four measures are of interest: how much time participants spend on background material to familiarize themselves with the scenario, how much time participants spent on the messages, how much time participants spent analyzing courses of action, and how much time participants spent generating or modifying options.

Time spent on background material was defined as time spent initially reading the three books provided to the participants, plus any time spent asking questions or discussing that material with the experimenters. Later reference to books could be situation assessment or often evaluation actively depending on the context of the action.

Time spent on the messages was defined as time spent reading and plotting (on the maps) the messages as well as time spent inferring the what the messages mean. The first two measures were directly observable from the participants' behavior. The last measure was derived from the think aloud behavior on the part of the participants and in response to direct probes from the experimenter.

Time spent on analysis was defined as time spent evaluating existing courses of action (including ones the participants themselves developed). Analysis was considered to have been performed when participants engaged in traditional analysis techniques such as wargaming, combat power analysis,

terrain analysis, task organization, decision analysis, and other analysis techniques specified by Army doctrine. The measure of analysis time was derived from two sources: the participants' verbal responses/think aloud and written analysis that could be directly observed.

Time spent on option generation was defined as time spent developing new courses of action or modifying old ones. This measure was derived from both verbal and written behavior on the part of the participants.

It should be noted that the measuring scheme used had certain drawbacks. Specifically, the assignment of participants' behavior to specific measurement categories necessarily contained a subjective element and attempts were made in design to minimize any subjective interpretations. For example, participants could be reading the messages and be generating a course of action at the same time. Unless participants articulated this or the experimenter was able to elicit this information with his queries, this activity would go undetected. Unfortunately, there appears to be no adequate solution to this problem. The position taken with respect to data analysis was that it is better to accept misses with respect to hypotheses tested than to allow false alarms. This was so because the experimental procedure had built in means for reducing misses (e.g., having the experimenter ask the participants questions), but none for safeguarding against false alarms when behaviors were "read" into what participants were doing. Therefore, unless there was observable evidence, as defined in the experimental design, that a behavior was occurring (e.g., a verbalization on the part of the participant or a written activity), it was assumed that the behavior was not occurring.

Also, the boundary between option generation and analysis was not always a sharp one. Given that for the present hypotheses, the distinction between option generation and option analysis is not crucial, the two measures were combined. Similarly, for background and message processing, the measures were combined as both involve getting familiar with the situation.

The second type of data that is of interest to the present study is whether participants located the 14th TA. This was important since it strongly influenced participants' perceptions of how favorable the tactical situation was. This measure was achieved in one of two ways. First, during the session, the experimenter looked for indications that the participants had found the 14th TA when they processed the messages. These indications came as participants discussed the messages or plotted them. Second, participants were explicitly asked where they thought the 14th TA was during the post-session interview.

The third type of data of interest to the present study is the participants' course of action recommendations. These data were collected from participants' course of action briefings at the end of the sessions. There were three basic types of courses of action available to participants: they could either withdraw, defend or attack with associated description detail for each, i.e., how it would be done. The first two would not accomplish the mission of achieving objectives along the IGB, but stood a better chance of preserving the fighting force. Only the attack option had the possibility of achieving the mission. Recall the earlier notions regarding a lexicographic ordering and lack of tradeoffs.

Results

No participants in this study chose to withdraw, the only options of relevance to the analyses below are defend and attack. Therefore, the courses of action were assigned to one of these two categories based on what participants stated in their course of action briefings.

Hypothesis 1:

"If participants think the best COA meets their target probability of success, they will focus on resolving uncertainties to verify its probability of success. In contrast, if they think that even the best COA has little chance of success, they will generate a new course of action."

This hypothesis speaks to the relationship between the probability of success of the best option and the aspiration level. According to our framework, if the best option appears to lie above the aspiration level but there is uncertainty in the situation, the decision maker will focus on resolving uncertainty in order to confirm that the best course of action will be successful. Conversely, if even the best course of action appears to fall below the aspiration level, then further situation assessment is likely to confirm this and is therefore not time well spent. Rather, the decision maker is likely to seek a new option and spend his time in option generation and analysis.

In order to test this hypothesis, a comparison is made of the time spent on background/situation assessment vs. analysis/option generation in both the Bad to Good and Good to Bad conditions. The prediction would be that in the Good to Bad condition, participants would spend more time on background/situation assessment compared to the Bad to Good condition and that the opposite would hold true for analysis/option generation.

This hypothesis does not distinguish between time stress and no stress conditions. However, since amount of time spent on each task correlated with the time condition, this hypothesis is tested using measures of percentage of time spent on each task. By using percentage of time, the time condition is controlled for without changing the intent of the hypothesis.

Table 3 below presents the mean time spent on background/situation assessment and analysis option generation in the four experimental conditions.

Table 3

Mean Percentage Time Spent on Background/Situation Assessment and Analysis/Option Generation for each of the Four Experimental Conditions:

	Background/Situation Assessment		Analysis/Option Generation	
	Bad to Good	Good to Bad	Bad to Good	Good to Bad
No Stress	81.43	90.23	18.57	9.77
Time Stress	75.56	90.23	24.44	9.77

As can be seen from Table 3, for background/situation assessment, Good to Bad participants spent on average 90.23% of their time on this, while Bad to Good participants spent on average 78.50% of their time. This difference was statistically significant, $F(1,28) = 4.42$, $p < .05$. Similarly, since analysis/option generation constituted the remainder of participants' time, the difference in percentage time here (9.77% and 21.50% for Good to Bad and Bad to Good participants, respectively) was also significant, $F(1,28) = 4.42$, $p < .05$. Thus, hypothesis 1 was supported.

Hypothesis 2:

"Time stress participants will focus relatively less time on situation assessment and relatively more time on option generation while no stress participants will do the opposite."

There are two tests of this hypothesis. First, one can look at the relative proportion of time spent by both time stress and no stress participants on each of these tasks. Second, if time stress participants are focusing relatively more attention on option generation than no stress participants, then time stress participants would be less likely to go with an existing course of action than no stress participants.

The first test of this hypothesis can be done using the data in Table 3. From Table 3, it can be seen that time stress participants, across conditions, spend on average 17.11% of their time on analysis/option generation while no stress participants spend on average 14.18% of their time on analysis/option generation. While this effect is in the predicted direction, it is not statistically significant, $F < 1$.

The second test of this hypothesis can be done using data on the number of time stress and no stress participants picking a new course of action versus an existing one. Table 4 below shows this data.

Table 4

Number of Participants Picking an Existing or New Course of Action by Time Stress and No Stress

	Existing	New
No Stress	5	11
Time Stress	1	15

As can be seen from Table 4, all but one time stress participant choose a new course of action while nearly a third of the no stress participants choose an existing course of action. Using Fisher's Exact statistical test for such contingencies with small cell counts, the probability of the results is .086. (This is the first of several tests using Fisher's Exact test. When the "probability of a result" is supported, this refers to the probability of the exact result obtained or one more extreme in the same direction, which is consistent with standard psychological significance testing procedures.) This is support, but the result is not significant at the conventional .05 level.

Hypothesis 3:

"When the situation initially looks good but then subsequent information suggests that the situation is bad, then time stress participants will be more likely than no stress participants to pick an existing course of action that would exceed the aspiration level in a good situation."

In order to test this hypothesis, the course of action choices are used in the Good to Bad condition. In this condition, three of eight no stress participants picked an existing course of action, while only one of eight time stress participants picked an existing course of action (See Table 5). This effect was actually in the opposite direction of the hypothesis, although the effect was not statistically significant. Using Fisher's Exact Test, the probability is .28.

Table 5

Choices of Courses of Action in the Good to Bad Condition

	Existing COA	New COA
No Stress	3	5
Time Stress	1	7

One possible explanation for this lack of effect may come from the hypothesis 2 results. There time stress participants were more likely in general to pick new courses of action than no stress participants. This effect may have overridden the hypothesized effect presented in hypothesis 3,

thus making time stress participants more likely than no stress participants to pick a new course of action instead of the other way around.

<u>Hypothesis 4:</u>

"When the situation initially looks bad but then subsequent information suggests that the situation is good, then no stress participants will be more likely than time stress participants to pick an existing course of action that does exceed the aspiration level."

In order to test this hypothesis, the course of action choices are used in the Bad to Good condition. In this condition, two of eight no stress participants picked an existing course of action, while no time stress participant picked an existing course of action (See Table 6). This effect was in the direction of the hypothesis but was not significant. (Fisher's Exact probability = .23)

Table 6

Choices of Course of Action in the Bad to Good Situation

	Existing COA	New COA
No Stress	2	6
Time Stress	0	8

The obtained result may have been mediated by the hypothesis 2 effect, i.e. that time stress participants were more likely in general to pick new courses of action than no stress participants. Since both hypotheses 3 and 4 are tied heavily to a perception on the part of participants as to what the situation is (as opposed to a general tendency for time stress participants to seek new courses of action), it is important to consider the course of action choices as a function of what participants' perception of the situation was. This can be done by looking at whether participants found the location of the 14th TA or not (participants' perception of the situation correlated perfectly with whether or not they found the 14th TA since those that did not assumed that the 14th TA would not play in their sector).

<u>Course of action selections as a function of whether or not participants located the 14th TA</u>

Analysis of the number of participants who located the 14th TA revealed a difference between the time stress vs. no stress conditions (see Table 7). In the time stress condition, only 3 of 16 participants located the 14th TA while in the no stress condition, 8 of 16 participants located the 14th TA. This difference was in the expected direction and was statistically significant, chi-square $(1, N = 32) = 3.46$, $p < .05$, one-tailed.

Table 7

Identifying 14th Tank Army (TA) Location as a Function or Time Stress versus No Stress

	Locate 14th TA	Did not Locate 14th TA	
No Stress	8	8	16
Time Stress	3	13	16

Earlier in testing Hypotheses 1-4, courses of action were characterized as old or existing versus new. Some of the participants who created new courses of action created defensive holding actions which had differing probabilities of mission outcomes. Examples include holding for awhile to await reinforcements with the hope of then continuing the attack. These actions can be generally characterized as attack or defend. While there was a significant difference in the number of time stress vs. no stress participants who located the 14th TA, the courses of action chosen by those who located vs. did not locate the 14th TA did not depend upon whether they were in the time stress or no stress condition. In other words, the variable of time stress appeared to have no affect on course of action selection other than to mediate the likelihood that participants would find the 14th TA. These results are displayed in Table 8. A summary 2 by 2 table showing course of action by time stress versus no stress shows no effect due to condition (chi square = 0.0) This coupled with Tables 5 and 6 confirm the stated result. As a result, further analyses are broken down by whether participants did or did not locate the 14th TA, rather than by time stress or no stress.

Table 8

Course of Action as a Function of Locating the 14th Tank Army and Time Stress/No Stress

		Locate 14th Tank Army	Did Not Locate 14th Tank Army
No Stress	Attack	5	7
	Defend	3	1
Time Stress	Attack	2	10
	Defend	1	3

	Attack	Defend
No Stress	12	4
Time Stress	12	4

Table 9 below shows the courses of action selected by participants as a function of situation (Good to Bad, Bad to Good) and by whether or not they located the 14th TA.

Table 9

Number of Participants Choosing to Defend or Attack as a Function of Situ..:ion and Whether they Located the 14th Tank Army

	Locate 14th TA		Not Locate 14th TA	
	Defend	Attack	Defend	Attack
Good to Bad	4	2	2	8
Bad to Good	0	5	2	9

A number of interesting findings come out of Table 9. First, when participants locate the 14th TA, the situation has a strong effect. Here, when the situation is bad (the 14th TA is present), two-thirds (four out of six) participants choose to defend. On the other hand, when the situation is good (the 14th TA is not present), all participants chose to attack. This effect is statistically significant, Fisher's Exact Test probability = .045. This result is really a test of the situation manipulation, i.e., does the presence of the 14th TA really change the situation from being good to being bad? Where the 14th TA was not located, condition make no difference (Good to Bad, Bad to Good), with Fisher's Exact probability = .581. Preference was to attack which was in line with stated mission orders.

What is particularly interesting here is that when the 14th TA is not located, initial situation has no effect. In both the Good-Bad and Bad- Good conditions, approximately 80% of the participants chose to attack. Even though participants in both conditions had different initial briefings as to where the 14th TA was expected to be, this briefing appeared to have no effect on the final recommendation (even though participants believed the initial briefing). This is constant with the prospect theory extension in which the participant near the aspiration level will prefer some chance of success over pure failure. Thus they would design actions that remove the negative potential outcomes of an attack.

Similarly, in the Good-Bad Condition, the initial briefing may suggest an outcome distribution that is also wide, but the bulk of it lying above the aspiration level. On the other hand, for participants who have located the 14th TA, the outcome distributions become much tighter given that the situation has less uncertainty in it. For the Bad-Good condition, the outcome distribution would be tight and above the aspiration level. For the Good-Bad condition, the outcome distribution would be tight and below the aspiration level.

Figure 6 illustrates how predominant course of action selected by participants relates to the outcome distributions suggested by situation condition and whether participants have located the 14th TA.

Figure 6 illustrates the treatment of uncertainty discussed above. When the situation is definitive, participants behavior polarizes to that suggested by that situation, i.e., participants in bad conditions defend while participants in good conditions attack. On the other hand, when the situation is uncertain, if it offers even the slightest chance for success, participants go with the best case assumption and attack. This is consistent with findings by Cohen, Tolcott and McIntyre (1987) who found that Air Force pilots would often ignore unconfirmed reports of enemy threats in a scenario setting. The rationale given by those pilots was that they could not allow themselves to be distracted from their missions unless they had concrete evidence (low uncertainty) that their mission was in jeopardy. The present participants showed this tendency as well.

The present findings raise the question as to whether uncertain negative outcomes are ignored in general or only when elimination of the uncertain outcomes would produce success. In other words, if the present scenario had been constructed such that even if the 14th TA did not enter the sector the mission would be in jeopardy would participants have attacked anyway? This would be analogous to a wide outcome distribution that lies entirely beneath the aspiration level. The question becomes whether participants in this case would defend rather than attack. This issue is important for the task of uncertainty management in decision aiding, but must be addressed in a later study.

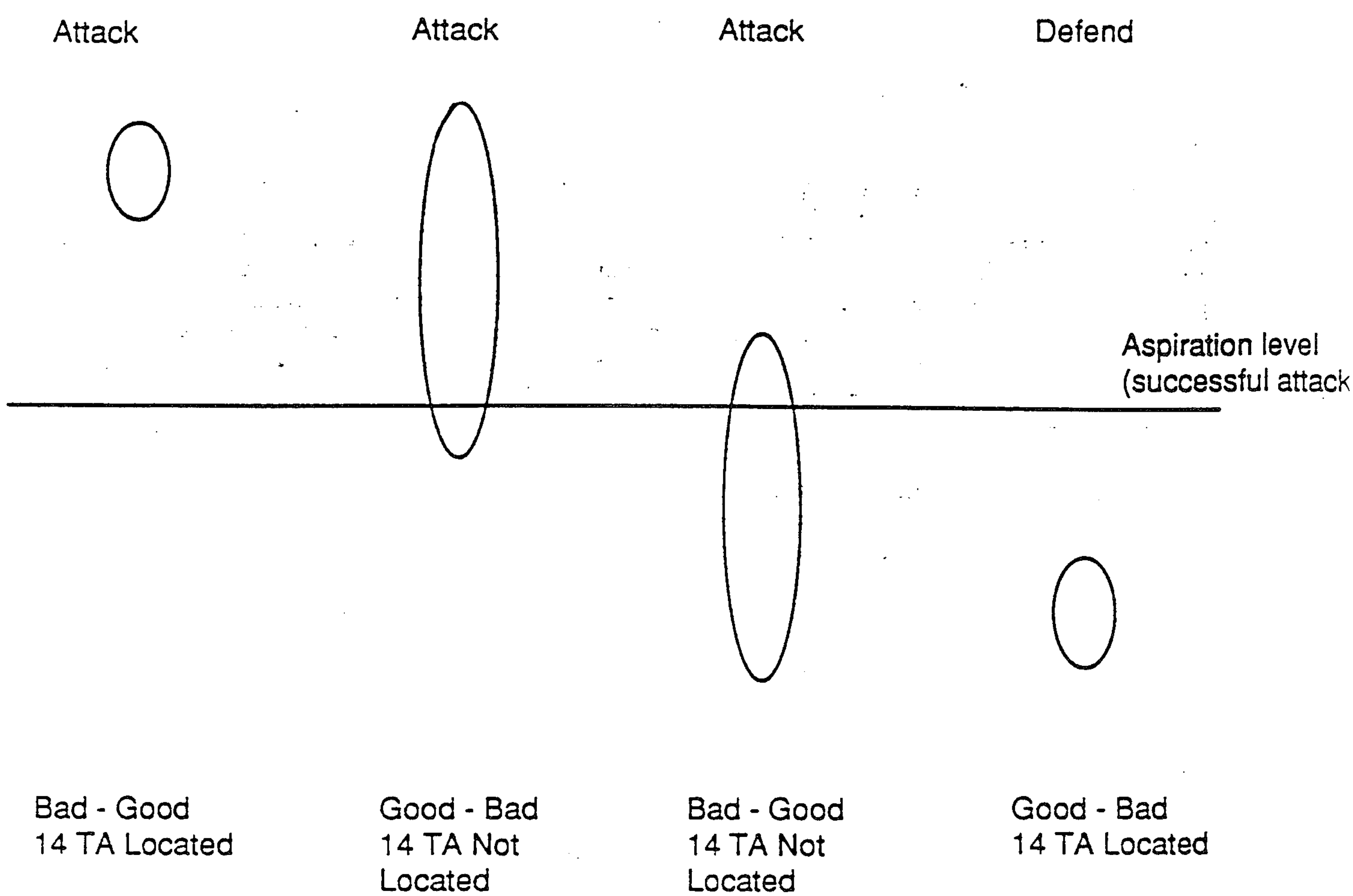


Figure 6. Outcome distribution suggested and predominant course of action by condition and whether subjects have located 14th tank army.

Aiding Concepts

As indicated, a post-experimental interview was conducted. One part of the interview requested aiding concepts that could have helped subjects with the experimental problem. Across 32 participants, 104 aids were listed (a participant could list more than one aid).

Twenty-one participants requested an intelligence aid to help them summarize information and depict the enemy situation. This was the single most requested aid and is consistent with the present contention that information management and uncertainty reduction is the single most important factor in course of action selection.

Twelve participants requested graphic/visual aids and templates to help present information in ways that could be abstracted quickly. This is also consistent with the above contention that information management is key in the decision making policy. This type of aiding was the second most requested.

The third most requested aid was a combat power and wargaming aid to help participants evaluate courses of action. Eight participants requested this type of aid. Collectively, these aids that were most requested by participants are consistent with the conclusions presented above.

Other aids requested by participants include: event templating, flow of battle depiction, aids that present status of assets controlled by the G-1, G-4, engineers and others, task organization aids, and milestone charting. Appendix A presents a summary of all the aids requested by participants.

The results of the course of action selection data and analysis of post-experimental question sessions, point to two important areas for aiding decision making, particularly under conditions of uncertainty and time stress. First, perceived combat power plays an important role in course of action selection. When participants were certain about the situation, their courses of action were driven almost entirely by combat power. Therefore, an aid that helps the G-3 and his staff compute and evaluate the effects of combat power in course of action wargaming should be useful to the G-3 and his staff.

Second, as the data suggest, combat power information is only influential when the information regarding it has a high degree of certainty. Therefore, an aid to help the G-2 and G-3 perform situation assessment to reduce and manage uncertainty appears to be the key to the course of action selection process.

DISCUSSION

The results of Study 2 show some support for the four hypotheses tested in the study and for the theoretical framework presented. Specifically, hypothesis 1 stated that participants whose initial impression of the situation and corresponding courses of action were good would spend more time on background and situation assessment processing and less in analysis and option generation than participants with the opposite initial impression of the situation. This hypothesis was supported. This suggests that participants did indeed try to reach their aspiration level in ways hypothesized by the proposed theoretical framework.

Hypotheses 2, 3, and 4 dealt more directly with the effects of time stress on decision making. Hypothesis 2 stated that time stress participants would focus more than no stress participants on generating new options and less on evaluating the situation as a means of solving a problem. This hypothesis was tested using two measures: time spent on these activities and actual course of action selections. Both tests were in the hypothesized direction but were not statistically significant.

One rationale for this hypothesis was that time stress participants might be more pessimistic about the situation than no stress participants. However, there was no evidence to support this explanation. In fact, the course of action recommendations when uncertainty was not resolved (the 14th TA not being located) suggested that participants in general were behaving as if they were using best case rather than worst case assumptions. While this data does not directly speak to hypothesis 2, its spirit is inconsistent with a pessimism explanation.

Hypotheses 3 and 4 stated that time stress participants would be less likely to pick up on changing situations and therefore be more prone to go with courses of action suggested by the initial situations. Hypothesis 3 addressed a situation that initially looks good but then shifts to bad, while hypothesis 4 addressed a situation that initially looks bad but then shifts to good.

Neither hypothesis 3 nor hypothesis 4 was supported. However, it is possible that the effects of both hypotheses 3 and 4 were overwhelmed by the effects of hypothesis 2 in that time stress participants in both situation conditions showed a tendency to go with a new course of action, regardless of the initial situation.

The most interesting results of the present study were not part of the original four hypotheses. First, while time stress did affect whether participants tended to pick a new course of action or stick with an old one, it did not appear to have a direct effect on the qualitative decision of whether to attack or defend. Rather the affects of time stress seemed to be mediated by whether participants did or did not find the 14th TA. Here time stress participants were less likely to find the 14th TA than were no stress participants, presumably because the latter had more time to process and correlate information.

When the 14th TA was located (and hence the bulk of the uncertainty associated with the situation resolved), participants' course of action recom-

mendations polarized depending upon the situation. When the 14th TA was not present, all participants chose to attack. When the 14th TA was present, two-thirds of the participants chose to defend, even though it ran counter to the commander's original mission. This finding did not interact with the time stress condition, suggesting that it is uncertainty resolution, not time stress per se that affects course of action recommendations.

When the 14th TA was not located (and hence uncertainty was not resolved), participants behaved virtually identically regardless of what initial information they were given about the situation. This is perhaps the most interesting finding of all. In essence, participants were not behaving as if they were discriminating between likely good outcomes and likely bad outcomes. This would suggest that participants were violating a standard Bayesian decision making approach, but has important implications for the practical aspects of military decision making and decision aiding.

As noted in the Results section, Cohen, Tolcott and McIntyre (1987) found that Air Force pilots tended to ignore unconfirmed reports of enemy threats. They felt that there was insufficient justification to abandon their mission and therefore it was not worth worrying about what they might have to face. Similarly, participants in the present study showed great reluctance to abandon what they perceived to be their mission. This supports the lexicographical nature of the tradeoff earlier discussed. It suggests that organizational decisions like these may be described by more complex, discontinuous rules.

These results suggest that participants require a high level of certainty before they are willing to adopt a conclusion to the extent that it will influence their decision. The problem arises when this high threshold results in the non-acceptance of a conclusion which is actually true. In the present study, of the ten participants who did not find the 14th TA when it was actually present in the participants' sector, eight chose to attack anyway. Hence, 80% of participants made the "wrong" (from the experimenter's viewpoint) decision in this condition. Two of eleven participants who did not locate the 14th TA when it was outside the participants' sector actually chose to defend rather than attack. Here, 18% made the "wrong" decision. All told, ten of twenty-one participants who did not find the 14th TA made a suboptimal decision--either failing to seize an opportunity that they had or perhaps even more costly, launching a highly risky attack against vastly superior forces.

These results have strong implications for aiding. First, to the extent possible, it is important to help decision makers resolve uncertainty. When decision makers have a clear and definitive picture, the present study suggests that they can plan effectively. However, such ideal conditions are rarely present. This is especially true in time stress situations where it is clear that participants will likely have to sacrifice uncertainty reduction.

In fact, the data underscore the need for aids to help process intelligence, especially under time stress, but even in cases where time stress does not exist. Recall from the results section that only three of sixteen participants in the time stress condition located the 14th TA and eight of sixteen participants in the no stress condition located it. On the other hand, participants in both conditions spent a large percentage of their time going through the messages that contained information regarding the location

of the 14th TA. Specifically, participants averaged approximately 45% of their time on the messages across both conditions.

In other words, going through the messages (i.e., processing the intelligence information) was a high cost, time-intensive task that produced relatively little payoff for a large percentage of participants, especially for time stress participants. This suggests a tremendous opportunity for aiding, either reducing the amount of time used to process intelligence information or increasing the effectiveness of that processing time.

Given that uncertainty was so frequently unresolved (based on the fact that 66% of the participants never located the 14th TA), it is important for decision makers to understand the potential impact of the uncertainty on their courses of action. Here, aids that not only present the different possible scenarios that decision makers might face in carrying out their courses of action, but also the impact of those courses of action on the success of their missions. However, the results of participants who did not locate the 14th TA suggest that simply having a qualitative understanding of whether the situation is predominantly good or predominantly bad is not sufficient (as participants in both conditions behaved identically). Perhaps a more graphic representation of the likely success or failure would be more appropriate so that participants could actually see the impact of these situations.

This discussion suggests aiding is needed in two areas. First, aiding is needed in uncertainty reduction and the presentation of situation information. This aiding would help decision makers understand what threat they are dealing with, which in turn gives them a sense of whether they can accomplish their mission or not. When outcomes depend on a critical event, methods for tracking the likelihood of that event in an efficient manner are highly desirable.

Second, aiding is needed that helps decision makers understand likely outcomes given different situation parameters. Such aiding might be in the form of wargaming tools that are tied to combat power (the results of the present study suggests that combat power plays a major role in driving Army tactical decisions).

These aiding suggestions are consistent with participants' own reports as to what types of aids they would like to have. Given that participants themselves actually request these types of aids, it is likely then that they would use such aids if they were available and implemented in a user friendly manner. Such aids would necessarily be useful for routine analyses, for no one will use an aid only in time stressed situations. Having decision makers actually use aids they are given is often an important issue in decision aiding. Decision makers are often presented with aids that are either too complicated to use or that the decision maker feels is not useful for their problem. The present study suggests that the two aids recommended here are both needed from an empirical perspective and desired by the decision makers.

A remaining empirical question is whether such aids would actually help decision makers in tactical situations. There are two related issues to consider--do they speed up or slow the decision making process and do they improve or degrade decision making quality? While the present study presents empirical support that aiding is needed, it does not address the practical

issue of whether implementing and using these aids will actually help or hurt. This is an important empirical question that will be addressed in Study 4.

Comparison of the Present Study with Previous Results

There was a noteworthy difference in the behavior of participants in the present study compared to that in previous studies (Leddo, Chinnis, Cohen and Marvin, 1987). In the present study, no stress participants spent approximately 80 - 90 minutes on coming up with a course of action recommendation, while in the Leddo et al. (1987) study, participants averaged about 180 minutes. The main discrepancy between the two studies appears to be in the amount of time that participants spent in analysis and option generation. In the present study, no stress participants averaged about 15 minutes on these tasks--in the previous study, it was much greater.

There are several possible explanations for this: different sample populations, different scenarios, different tasks. In the previous study, participants were all lieutenant colonels who were instructors at the Combined Arms Services and Staff School (CAS3). The present study had participants of different ranks (captain to lieutenant colonel) and different assignments. While the populations across the two studies were disjoint, the diversity of participants used in the present study, coupled with the vast discrepancy in behavior suggests that sample characteristics may not alone have caused this difference.

A likely explanation lies with the nature of the task. The previous study scenario used a defense problem while the present study used an offense one. There is no reason to think that this variable alone would cause such a great effect, unless one problem was inherently more difficult than the other. The previous study problem was rather stereotypic of TRADOC problems and most participants there felt it was not too difficult. The present problem was more novel and participants felt it was challenging. This would seem to require more rather than less analysis and work against the findings here. However, the more novel problem may not lend itself to the standard analytic procedures taught in the different military schools. Such procedures can be quite time consuming. Thus, the defensive problem that allowed use of such time consuming procedures could be expected to result in more analyses than the novel offensive problem and this was the result. Assuming that participants do not simply do analyses in a ritualistic procedural manner, the analyses have value. If the novel situation does not lend itself to standard analytical procedures, this suggests that an aid could provide a means to address such "novel" situations.

The tasks presented to the two sets of participants were different in a subtle, but perhaps powerful way. In the previous study, participants were given three defensive courses of action and asked to make a recommendation. In the present study, participants were given a withdraw and an attack course of action and given the same instructions. In the first study, the commander appeared to have already set the division's policy, i.e., that the division would defend. The participant was then left with the task of making a fine grain analysis of which course of action would then implement this policy. Such a task may have required an in-depth analysis, and the previous study results should that participants tended to use several analytical methods to arrive at this decision.

In the present study, the policy had not yet been worked out. Rather, participants had to decide the more basic issue of whether to attack or defend rather than how to attack (beyond a general description). This policy decision seemed to be fundamentally different than the implementation decision made by participants in the previous study. The policy decision seemed to rely more on assessing what the situation warranted rather than trying to figure out how to make something work. In fact, participants in the present study spent approximately 85% of their time forming an impression of the situation, which was far greater than the percentage of time so spent by the previous study participants.

Participants in the present study also shied away from the more formal analytic techniques such as combat power analysis that participants in the Phase I study used. The present participants wanted a subordinate to perform these analyses, almost suggesting that in their policy position, they did not feel it was their responsibility to perform them.

What is also interesting is that participants in the present study appeared to use different decision criteria than those used by participants in the previous study. In that study, participants often cited criteria such as protecting the rear area and providing sufficient reserve as criteria driving their decisions. Such criteria appear to be focused on successfully implementing a particular course of action. In the present study, participants cited decision criteria such as not exposing the flank of the corps main attack and preserving the fighting force for follow-on missions. Such criteria are reminiscent of the policy-type decisions one would expect a commander to make.

The most interesting difference between the previous and present studies was the amount of time taken by participants and the implications it has for aiding, particularly under time stress. As noted above, participants in the present study took approximately half as much time to recommend a course of action as did participants in the previous study. This suggests that policy decisions are made much faster than implementation decisions. We would expect that policy decisions would be much more robust under conditions of time stress than would be implementation decisions. This suggests that aiding attempts would best be focused on course of action implementation rather than policy analysis.

Also, the types of aiding that one would give to policy makers would be different than that given to course of action planners. Policy makers have a great need for information regarding the situation (as evidenced by the amount of time the spend on situation analysis). Therefore, aids which help them process large quantities of information rapidly and accurately would be of great help. Policy makers also spend little time looking at course of action detail. However, since participants in the present study stated that they wanted other staff members to perform these functions, it is probably unwise to force policy decision makers to conduct detailed decision making. Policy makers do need aiding in terms of depth of decision making, but this aiding should be directed at providing staff members who perform these functions. Automated aids would then be directed at these personnel.

STUDY 3

INTRODUCTION

Goal

The goal of the Study 3 was to gain feedback regarding potential aiding concepts for decision making under time stress and uncertainty that would be later used in an experimental setting to determine whether implementation of those aiding concepts would lead to better and faster decision making.

Background

Study 2 identified two prominent areas that aiding appeared most promising: presenting current intelligence in a summarized manner, and course of action analysis that centered on representing relative combat power between friendly and enemy forces.

In Study 3 displays representing potential aids and how users might interact with them were constructed. The specific topic areas covered were: strength of enemy units, combat power ratios, categorizing units by common parents, representation of uncertainty, identifying potential threats to missions, decision templating, and course of action selection and analysis.

METHOD

Participants.

Twenty participants in all were tested. All participants were either captains or majors and had a variety of backgrounds (staff members, combat support, artillery; division, brigade and battalion experience). Most participants had been to CAS3 and some had also been to the Command and General Staff College at Ft. Leavenworth (CGSC).

Materials.

The materials used were sample paper displays of the potential aids described above. The aiding topics were ordered sequentially with subsequent displays building on prior displays. The aids used and order followed were strength of enemy units, combat power ratios, categorizing units by common parents, representation of uncertainty, identifying potential threats to missions, decision templating, and course of action selection and analysis.

Procedure.

Participants were tested individually, except that the last three participants were tested together due to scheduling irregularities.

Each session consisted of three phases which occurred as iterative cycles for each topic area covered. First, participants were asked a series of prepared questions regarding each topic. This was done to provide additional background to the topic areas and help identify important issues that might help refine the design of the aiding concepts.

Next, participants were shown sample aids for the current topic. Typically, participants were provided several alternative ways of presenting the information depicted in the display. They were asked to state their preferences (in rank order) and comment on what they liked and did not like about the displays. Participants were also asked to make suggestions regarding how the displays could be improved.

Third, for the topics of enemy strength, combat power, and course of action selection, participants were asked to make a series of judgments using the displays. For enemy strength and combat power, participants made judgments regarding combat power using the different display formats. For course of action analysis, participants were presented with two stylized versions of the present study scenario. One scenario depicted the 14 Tank Army in the friendly sector, the other depicted the Army out of the sector (thus these scenarios depicted the two versions in Study 2, with the intelligence information already processed). As in the Study 2, participants were asked to recommend a course of action.

In cases where multiple judgments were required regarding a topic area, the materials presented were in counterbalanced order across participants.

RESULTS

Enemy Strength

There were six variations of symbology offered to the participants in order to reflect the strength of enemy units (See Figure 7). The first was the standard symbol currently used by the Army that does not display enemy strength (Not shown in Figure 7). The second method used the standard symbol but shaded it in proportionate to the strength of the enemy unit (e.g., a unit at 50% strength would be shaded in half way).

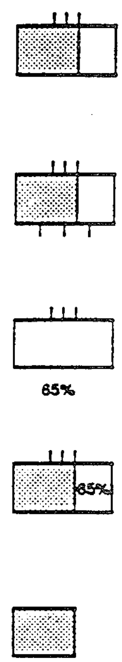


Figure 7 Unit strength variations.

The third method was similar to the second except that the bottom of the symbol had quartiles marked off (25%, 50%, 75%) so that participants could have some gauge as to how far the unit was shaded in. The fourth method had the standard unit symbol but underneath the symbol was a written expression of the enemy strength. The fifth method was a combination of the second and fourth methods using both shading and numerics to reflect enemy strength. The sixth method used shading, although this time the unshaded portion of the symbol was eliminated. Therefore, the symbol size reflected the strength of the unit.

Based on this description, the second, third and sixth methods conveyed unit strength visually (by shading), method four conveyed unit strength numerically, and method five used a combination of visual and numeric methods.

Of twenty participants surveyed, 9 participants preferred the combination numeric and visual representation, 5 preferred shading alone (without

markers below the symbol), 4 preferred numerical representations, and 2 preferred the standard.

Those hesitant to use the new symbols were afraid of the additional work that providing such information on the maps might create. Those that preferred visual shading (both alone and in combination with numerics) liked the idea that one could get a quick glance at the map and see the strengths portrayed visually. Those preferring numerics liked the precision the numerics provided.

Participants were given examples of units whose strengths were portrayed using methods 2,3 and 6 (See Figure 8). Their task was to estimate the strength of the units portrayed. Of interest is their accuracy. Since the standard method does not convey strength information, there was no need to use such examples. Methods 4 and 5 which use numeric representations of strength were not used since inferring strength from numbers is a trivial process.

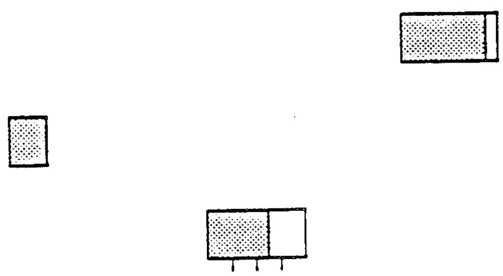


Figure 8 Estimate strength of each unit.

Three examples from each method were used. Units were selected at 37.5%, 62.5%, and 87.5% strength. (These figures were partly driven by the graphics package used to generate the symbols--it could delineate up to eight-inch increments.) Judgment results are described below in Table 10.

Table 10

Judgments of Enemy Strength Using Different Representations

Strength:	37.5			62.5			87.5		
Symbol Method:	2	3	6	2	3	6	2	3	6
Estimate Range:	30-45	33-45	25-40	60-75	60-70	40-70	80-90	85-90	75-100
Greatest Error:	7.5	7.5	12.5	12.5	7.5	22.5	7.5	2.5	12.5
Modal Response:	35	35	30	65	60	50	90	85	100
Modal Error:	2.5	2.5	2.5, 7.5	2.5	2.5	12.5	2.5	2.5	12.5
Avg. Error:	4.7	3.0	6.3	3.1	2.6	10.8	2.7	2.2	8.8
% Avg. Error:	12.5	8.0	16.8	5.0	4.0	17.3	3.1	2.5	10.1

As can be seen from the above table, participants are generally fairly accurate in judging enemy strengths. The most difficult symbol type was the shaded symbol whose size was proportional to its strength. Therefore, participants appeared to need the gauge of the unshaded portion of the box to help with the judgment.

On the whole, participants generally were able to estimate enemy strength using shading to within 2.5 percentage points of the actual value. Therefore, the percent error was directly influenced by the level of the strength of the enemy (e.g., 2.5 is a larger percentage of 37.5 than it is of 87.5). Although having the quartile markings did appear to improve accuracy slightly, participants were generally firm about not wanting them. However, the combination of visual and numerical representations (method 5) of enemy strength should solve any accuracy problem with the visual estimating process alone.

Combat Power

There were three variations of combat power representations presented (See Figure 9a-c). The standard approach was to present participants with the standard symbols and allow them to use order of battle and other information to calculate combat power ratios (Figure 9a). The "visual" method, varied the size of the symbol in proportion to its combat power (Figure 9b). For example, the symbol for a division is larger than that of a brigade. The symbol is then shaded in direct proportion to its percent strength (see method

2 above). Therefore the total area that is shaded would visually depict the combat power of the unit. A comparison across friendly and enemy units could be made by visually comparing the amount of shading on both sides of the FEBA. The third method was numeric (Figure 9c). Here the standard symbol was used, but a measure of combat power is placed below the symbol (much like method 4 in the enemy strength judgment).

Participants were instructed, for both the visual and numeric methods, to assume that there existed a formula that satisfactorily determined what the combat power of the units was (i.e., that reliably produced the combat power number or unit size for the numeric and visual representations, respectively). Participants went along with this assumption but acknowledged the importance of understanding where those numbers (or sizes) came from in order to feel confident using them.

Of the representation schemes presented, 61% preferred the numeric representation as their preference, 22% preferred the visual, and 17% preferred the standard representation (which did not depict combat power). Three participants explicitly requested a combination of the visual and numeric techniques. The visual representation was also the one that was most often ranked least preferred (by 77% of those making such a ranking). The concern most often expressed with the visual representation was that participants feared that accuracy would be a problem for manually drawn symbols (i.e., in the field, soldiers tend to be casual about drawing symbols and the level of precision required would be too difficult). As a result, the visual symbol may be appropriate only for automated graphics.

For the series of judgments, participants were provided with two sets of combat power ratios to judge. These judgments depicted a friendly division opposing an enemy Army. Each set contained one standard judgment where participants were required to use order of battle and staff estimate information to compute combat power ratios, two visual representations where participants compared the relative size of shaded areas, and two numerical representations where participants essentially performed arithmetic to judge relative combat power. The first set contained "disaggregated" units, i.e., the division was broken down into its brigades and the enemy Army was broken down into divisions. The second set aggregated the friendly and enemy units so that the participant was comparing a whole Army versus a whole division. It was hypothesized that the participants would be faster and more accurate in judging the aggregated examples than they would the disaggregated ones. It was also hypothesized that participants would be faster in making judgments using the visual and numeric methods than they would using the standard methods.

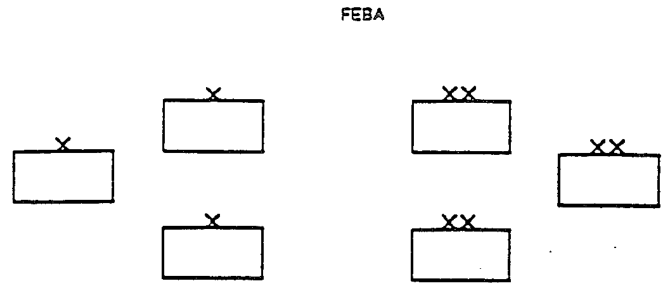


Figure 9a. Relative combat power judgment using traditional method (with OB charts)

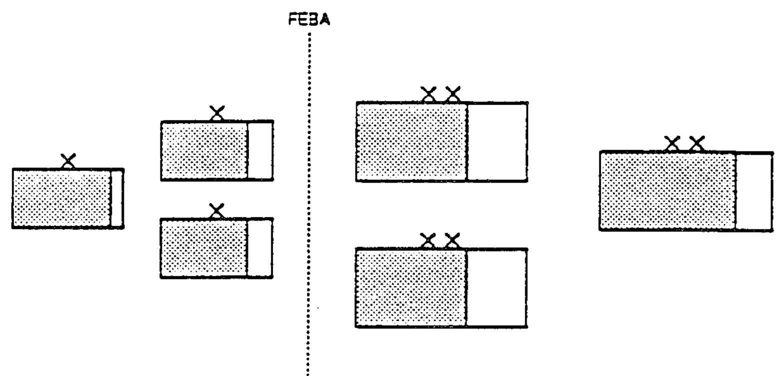


Figure 9b. Relative combat power using visual judgment.

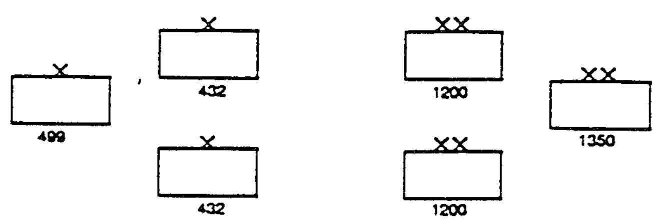


Figure 9c. Relative combat power using numerical judgment

Accuracy of estimate could be determined for both the visual and numeric examples. However, different methods can be used to compute combat power the "standard" way. Therefore, accuracy cannot be judged since the method used by participants could not be determined. However, the range of the estimates provided is interesting to report since it suggests that different staff members might be operating under different assumptions while looking at the same situation. For example, in the standard representation mode, disaggregated example, of the 16 participants giving a response (four participants said that they did not know how to compute combat power ratios), 6 participants thought the friendly forces were stronger, 3 judged them equal, and 7 judged the enemy forces stronger. If nothing else, aiding might be a useful tool to gain consistency (if not accuracy as well).

Table 11 below presents data regarding estimates (and accuracy where measurable) and response times:

Table 11

Disaggregated and Aggregated Judgments of Relative Combat Power Using Different Representation Methods

	Disaggregated Examples				
	Standard	Visual1	Visual2	Numeric1	Numeric2
actual force ratio:	XXXXX	.58	.50	.36	.53
estimate range:	.2 - 3	.13 - 3	.11 - .83	.29 - .5	.33 - .8
(w/o outliers):	.2 - 3	.29 - .67	.25 - .83	.29 - .5	.33 - .8
avg. error:	XXXXX	35.5%	38.2%	10.7%	10.5%
average response times:	7:35	1:31	:43	:46	:40

	Aggregated Examples				
	Standard	Visual1	Visual2	Numeric1	Numeric2
actual force ratio:	XXXXXX	.51	.37	.36	.53
estimate range:	.25 - 1.5	.25 -.67	.25 - .46	.30 - .40	.33 - .62
avg. error:	XXXXXX	26.2%	30.0%	6.80%	7.50%
average response times:	4:11	:40	:24	:19	:12

There are several trends worth noting here. First, judgments of aggregated information are faster and more accurate (where measurable) than judgments based on disaggregated information. Both the visual and numeric schemes produced faster judgment times and narrower estimate ranges (i.e., there was more consensus on the force ratios than in the standard condition).

Looking at the ranges proves interesting. In the disaggregated, standard example (which is how officers would normally get the information), the range of .2 to 3 represents a factor of 15 in terms of discrepant view of the battlefield. Nor could .2 and 3 be considered outliers as another participant gave an estimate of .21, while one gave an estimate of 2.6. This factor of 15 has important implications. In essence, participants' estimates ranged from the enemy having a 5:1 advantage to the friendly forces having a 3:1 advantage. Normally, when the enemy has a 5:1 advantage, the friendly forces defend, whereas when the friendly forces have a 3:1 advantage, they attack. In other words, the discrepancies in the participants' responses were great enough to suggest that diametrically opposed courses of action would be recommended based on different perceptions of the same situation. This discrepancy suggests the need for aiding to gain a greater consistency in the estimation process. Even if accuracy is not obtained, consistency will at least help insure that the different combat elements will be working on the same problem.

While one of the disaggregated visual examples produced a range of .13 to 3, which is an even greater spread, both of these limits seem to be outliers since the next closest estimate on each side of the range is different by a factor of at least two. With this revised range, of .29 to .67, the spread is now a factor of 2.3. The second example produced a range of .25 - .83 (again there was one outlier who gave an answer of .11 - he was also the one who produced the outlier of .13 for the other example) or a factor of 3.3. While this range may also be unacceptable, it is certainly much lower than the range for the standard example.

The numeric examples produced a somewhat narrower range than the visual examples. Example one had a range of .29 - .5 with a factor of only 1.7, but the second example had a range of .33 to .8 with a factor of 2.4.

It is interesting to note that with the exception of one outlier for one of the visual examples, for both the visual and numeric representations, all participants correctly perceived that the enemy force was larger than the friendly force (i.e., for 80 total judgments, 99% of them had the correct qualitative judgment that the enemy was larger). Such a degree of consensus was not found in the standard example, where nearly as many participants thought the friendly force was larger as thought the enemy force was.

The average error for the numeric examples was also lower than for the visual ones (e.g., 10.6% versus 36.9%, respectively). Similarly, response times for the numeric examples were the fastest, although participants should substantial improvement in response times between their first visual example and their second, suggesting a practice effect. Given that participants probably have more experience using arithmetic than working with visual judgments such as these, it would be interesting to see how they would perform over extensive trials.

The aggregated examples saw similar patterns to the disaggregated ones, however, all examples showed improved speed and accuracy. In particular, the numeric examples saw speed cut more in half as participants were able to compare two numbers versus two groups of three numbers each. In addition, accuracy was improved so that average error was 7.2%.

The aggregation also produced more consistency in the estimation process. For the standard example, the range was .25 to 1.5 for a factor of 6. While this may be disturbing, it is much better than the previous factor of 15.

For the aggregated visual examples, example one had a range of .25 to .67 or a factor of 2.7 while the other example had a range of .25 to .46 or a factor of 1.8. For the numeric examples, the first had a range of .3 to .4 for a factor of 1.33, while the second had a range of .33 to .62 or a factor of 1.9.

In conclusion, the numeric representation seemed to be the best all around means to express relative combat power, particularly if forces can be aggregated. The numeric representation was quickest and most accurate, it built the most consensus and was the one preferred most by participants. On the other hand, the standard method was by far the slowest and produced the least consensus.

Classification of Units

Normally when maps are posted, units are designated by their identification number (e.g., the 1st brigade of the 16th division). Boundaries between units are drawn to identify which units go together and to give insight into the command and control parameters of the unit. In the present study scenario, friendly and enemy forces were depicted as intermeshing. Therefore, it would be more difficult to group forces.

Alternatives to this standard scheme would be to color-code units based on their parent. For example, all regiments belonging to a common division would be color-coded one way, while those belonging to another division would have a different color. There are two variations on the color code theme--different units could be color-coded using different shades of the same color (e.g., enemy units could be different shades of red), or different units could have discreet colors (e.g., green vs. yellow vs. blue). It should be noted that these methods also use the standard unit designators and boundary markings, so that the only change to the standard symbology would be color-coding.

Participants were asked to state their preference for the standard coding scheme or for color-coding. Those that preferred color-coding were then asked whether they would prefer color coding using shading or discrete colors.

Of the 20 participants surveyed, 9 preferred the standard method, while 11 preferred color-coding. Of those 11, 8 preferred discrete colors while 3 preferred different shadings.

Potential problems with the color-coding that participants cited are that some people are color-blind, and that maps are often read under different lighting conditions--hence it may be difficult to pick out colors. On the

other hand, people could simply ignore the colors under these conditions and use the unit designator markings and boundaries.

Uncertainty

An important topic is uncertainty. We have been dealing with "firm" estimates of strength and combat power ratios. In the field, such precise information is often not available. However, in the field, uncertainty is often ignored or simply alluded to in briefings (per participants' comments). We explored several alternatives for representing uncertainty regarding enemy strength.

The standard unit symbol typically does not express uncertainty except in cases where the symbol is drawn with a dotted line to communicate that the G-2 has inferred that the unit is there (e.g., from doctrinal templates), or the G-2 places a question mark near the symbol to indicate information is missing.

Six alternative methods for representing uncertainty were also presented (See Figure 10). Method 1 (called the "interval" method) depicted enemy strength as an interval (e.g., between 50% and 75%). Graphically this was depicted by having the certain portion of the estimate (50% and below) darkly shaded and the uncertain portion (50-75%) lightly shaded.

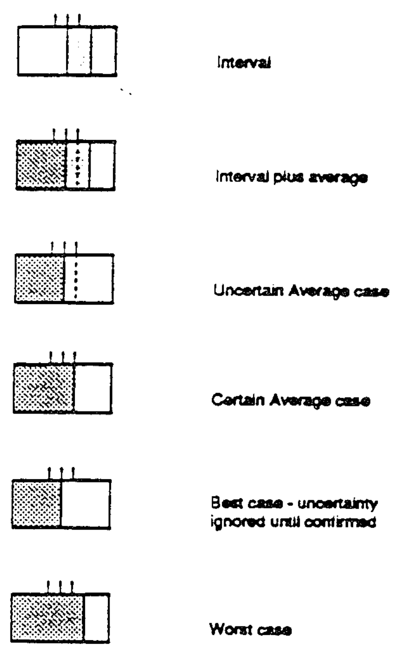


Figure 10. Uncertainty regarding enemy strength.

Method 2 (called the "interval plus average" method) used the method one symbol, plus drew a dotted line through the middle of the interval to indicate an estimate of the most likely estimate of enemy strength.

Method 3 (called the "uncertain average") used the method 2 symbol except the interval was eliminated (in other words, the certain 50% is presented with dark shading, plus a dotted line is draw to show that the strength is estimated to be higher, e.g., 62.5%).

Method 4 (called the "certain average") had its symbol darkly shaded in to the point of the uncertain average's dotted line. In other words, this method treated the average estimate as a certain one. Hence, this symbol does not really convey uncertainty.

Method 5 ("called the "best case") was similar to method 4, except that the symbol was darkly shaded only as far as the bottom portion of the method one interval (i.e., 50%). In other words, this method assumes that unless confirmation is received that the threat really is worse than the confirmed 50% strength, the assumption would be that the strength was at 50% or the best case. This method also does not convey uncertainty.

Method 6 (called the "worst case") was similar to method 5, except that the symbol was darkly shaded to the up limit of the method one interval (i.e., 75%). This method assumes worst case and plans for that. It also does not convey uncertainty.

Of the 20 participants surveyed, the results for the number of participants who picked each method as 1st and 2nd preference are presented below:

1st choice	2nd choice
interval (7)	interval (5)
worst case (7)	certain average (3)
uncertain avg. (3)	interval plus average (2)
interval + avg. (2)	best case (2)
certain avg. (1)	worst case (1)

As can be seen, there were no clear winners. However, 60% of the participants preferred that some type of uncertainty be represented (interval, uncertain average and interval plus average). Many participants stated that they had seen uncertainty described in each of the six ways presented to them (although not necessarily using our symbols). Virtually all participants had worked with intervals, averages, and worst case methods of addressing uncertainty.

Interestingly, interviews with participants gave divergent reasons for their preferences. While many participants wanted to plan for worst case scenarios, others felt that always going with worst case scenarios could be stifling. Participants who liked the interval methods claimed that it enabled them to see best case, worst case, and in-between.

It is unclear what the implications of this survey are. While the preponderance of participants did want to see uncertainty represented, a sizable portion of participants did want to be able to plan for worst case. While the interval representation does allow this, it may be useful to give decision makers the options of working with a variety of assumptions (e.g.,

best, average, or worst case) but keeping the uncertainties and ranges of possible outcomes salient.

Participants did seem highly attuned to the "fog of war" or uncertainties inherent in the battlefield. Therefore, it is probably wise to have aiding concepts which aid uncertainty management and planning. In other words, participants appear to want to avoid being stifled by uncertainty, hence any aid should probably help them plan in light of uncertainty, rather than "rub the uncertainty in their faces" and stifle them.

Intelligence Aid for Enemy Threat

Study 2 results suggest, and Study 3 supports, the notion that managing time stress and uncertainty revolves heavily around information flow and management. There was a strong consensus in Study 3 that intelligence information needs to be disseminated and processed better. Specifically, subordinate units and cell at the same level (e.g., G-3, G-4) need current intelligence information. Also, these additional cells need the information in summary and categorized form so they can pick and choose what they need rather than go through a series of reports and messages to decide what is relevant and what is not.

In this spirit, screens were presented from a sample intelligence aid that could be disseminated to other cells in the decision making network (See Figure 11 a-d). This aid would allow the decision makers to access information in a variety of ways and levels of detail (e.g., by unit, sector, time) and call up what they need at the level of detail they wanted it.

Response to the displays was very positive. Participants found all the information valuable and particularly liked the aspects of the aid that helped draw conclusions about enemy activities and intentions. However, participants showed strong consensus in wanting the G-2 to put his stamp on the aid's conclusions, i.e., they did not have faith that the aid could process information and infer enemy intentions. Therefore, they viewed the tool as a database management tool, rather than as an expert system.

Potential Threats to Mission

Given that participants are particularly interested in inferential aspects of intelligence, i.e., the impacts of the enemy situation on potential friendly courses of action, a series of displays were developed that depict potential threats to a given mission/course of action. The mission selected was the present study scenario depicting a planned division attack to seize two objectives. The displays depicted potential enemy threats with arrows drawn to the places along the avenues of approach that they were expected to engage the friendly forces.

OPTIONS: OB PREVIOUS ACTIVITY SIMILAR ACTIVITY INFERENCE

- * ID
- * Parent
- Subordinate
- * Strength
- Disposition
- * Current Activity

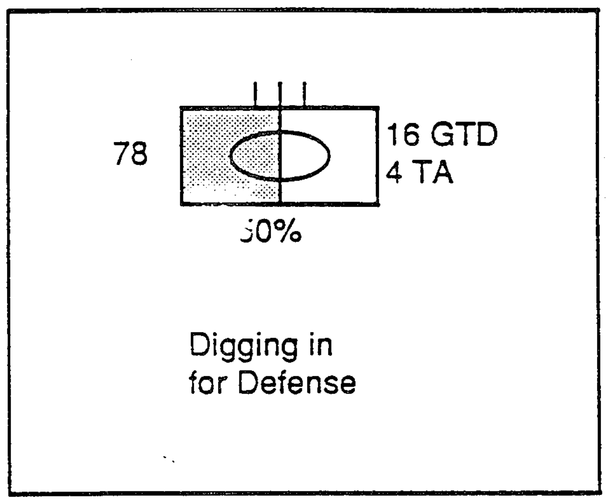
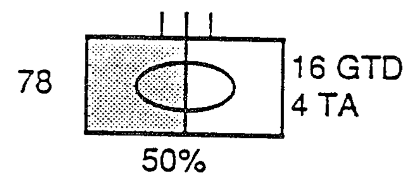


Figure 11a Intelligence aid for enemy threat: OB displayed.

OPTIONS: [OB] [PREVIOUS ACTIVITY] SIMILAR ACTIVITY INFERENCE

OB	PREVIOUS ACTIVITY
* ID	* 0600
* Parent	* 0500
Subordinate	* 0400
* Strength	0300
Disposition	
* Current Activity	



Digging in for Defense 0600

Digging in for Defense 0500

Limited Attack 0400

Figure 11b Intelligence aid for enemy threat: OB, previous activity displayed.

OPTIONS: OB PREVIOUS ACTIVITY SIMILAR ACTIVITY INFERENCE

* ID
Parent
* Strength
Disposition
* Current Activity

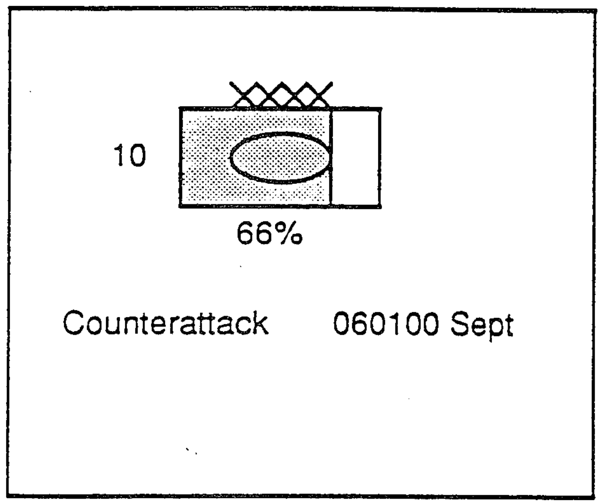


Figure 11c Intelligence aid for enemy threat: OB, similar activity displayed.

OPTIONS: OB | PREVIOUS ACTIVITY | SIMILAR ACTIVITY | INFERENCE

* ID
  Parent
* Strength
  Disposition
* Current Activity

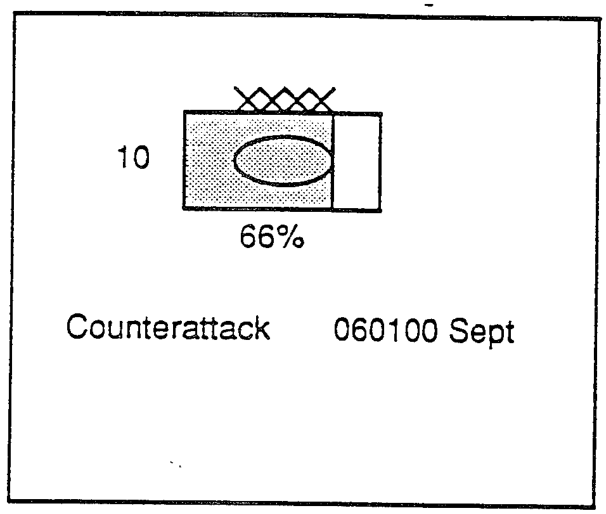


Figure 11d Intelligence aid for enemy threat: OB, similar activity and inference displayed.

There were four variations used (See Figure 12 a-d). The first used standard symbology, with the arrows drawn. The second used the standard symbology with unit strength presented below it. The third used the visual method of expressing both friendly and enemy combat power (see combat power section). The fourth used both the visual method of expressing combat power and the unit strength below it.

Again, participants were asked to rank order their preferences. Table 12 below lists the frequency with which each representation was ranked 1st, 2nd, 3rd, and 4th, as well as mean rankings.

Table 12

Mean Rankings of Different Representations of Potential Threats to Mission

	standard	strength	combat power	combat power + strength
rank 1st)	0	3	6	9
rank 2nd)	2	3	6	6
rank 3rd)	1	11	2	3
rank 4th)	14	0	3	0
mean ranking	3.71	2.47	2.12	1.67

As can be seen from Table 12, the combat power and strength format was favored most, followed by the combat power alone. Typically, the participants' reasons behind the preference was that these displays conveyed the most information. Interestingly, a display that had the visual representation of combat power (either alone or in conjunction with strength), was ranked first over 83% of the time by participants. This is interesting in light of the fact that visual representation of combat power as a stand-alone symbol was not very appealing to participants. However, when integrated in a specific application where participants could get a quick visual representation of a likely encounter and the relative power of the forces involved, this format was highly preferred. This suggests that the appeal for different types of information representations may depend on the applications that they are being used for (which should not be a counter intuitive conclusion).

In addition to the visual depiction of the battlefield and potential threats, participants were particularly interested in seeing time lines of how events would unfold. This could be done in a variety of ways. First, rates of movement could be provided as well as time frames for specified actions. Second, participants could be provided with different time slices so that they could see how events would evolve. Third, with the aid of computer animation, participants could be presented with moving pictures (e.g., as in a movie or cartoon) that depict how events will flow on the battlefield.

Decision Templating

The potential threats to mission concept was integrated with the decision template that is used as part of the intelligence preparation of the

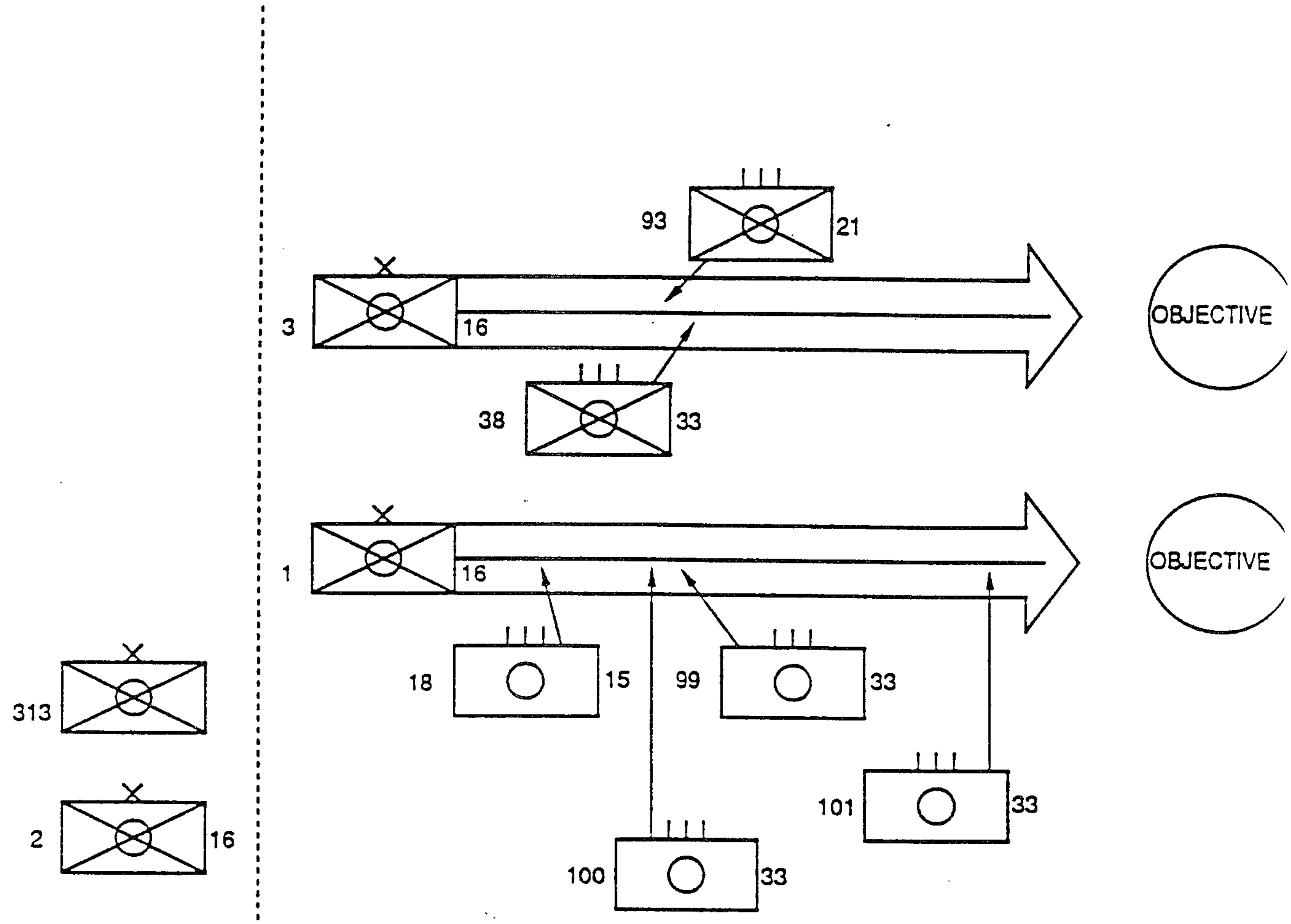


Figure 12a. Potential threats to mission: Standard symbology

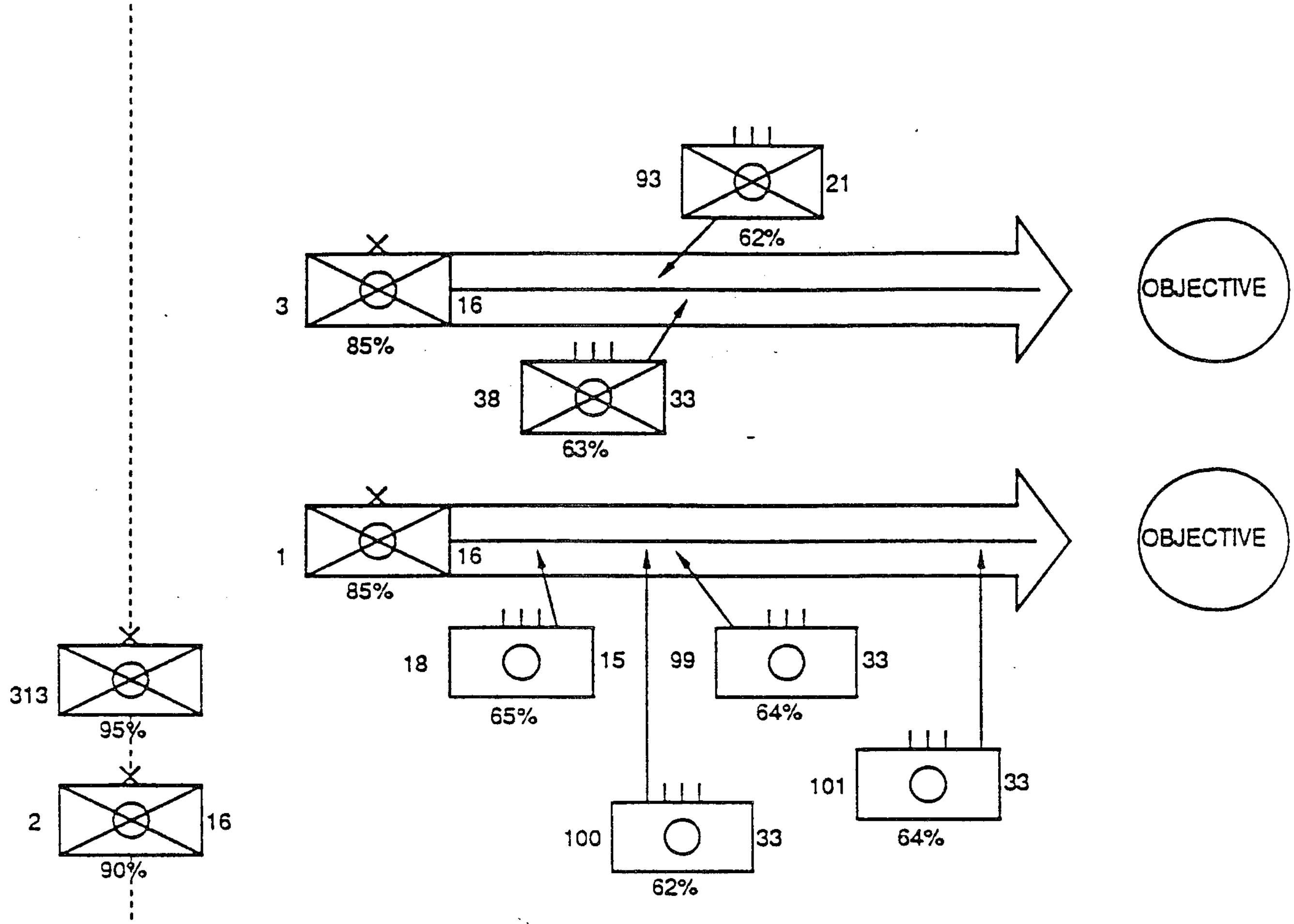


Figure 12b. Potential threats to mission: Unit strength version

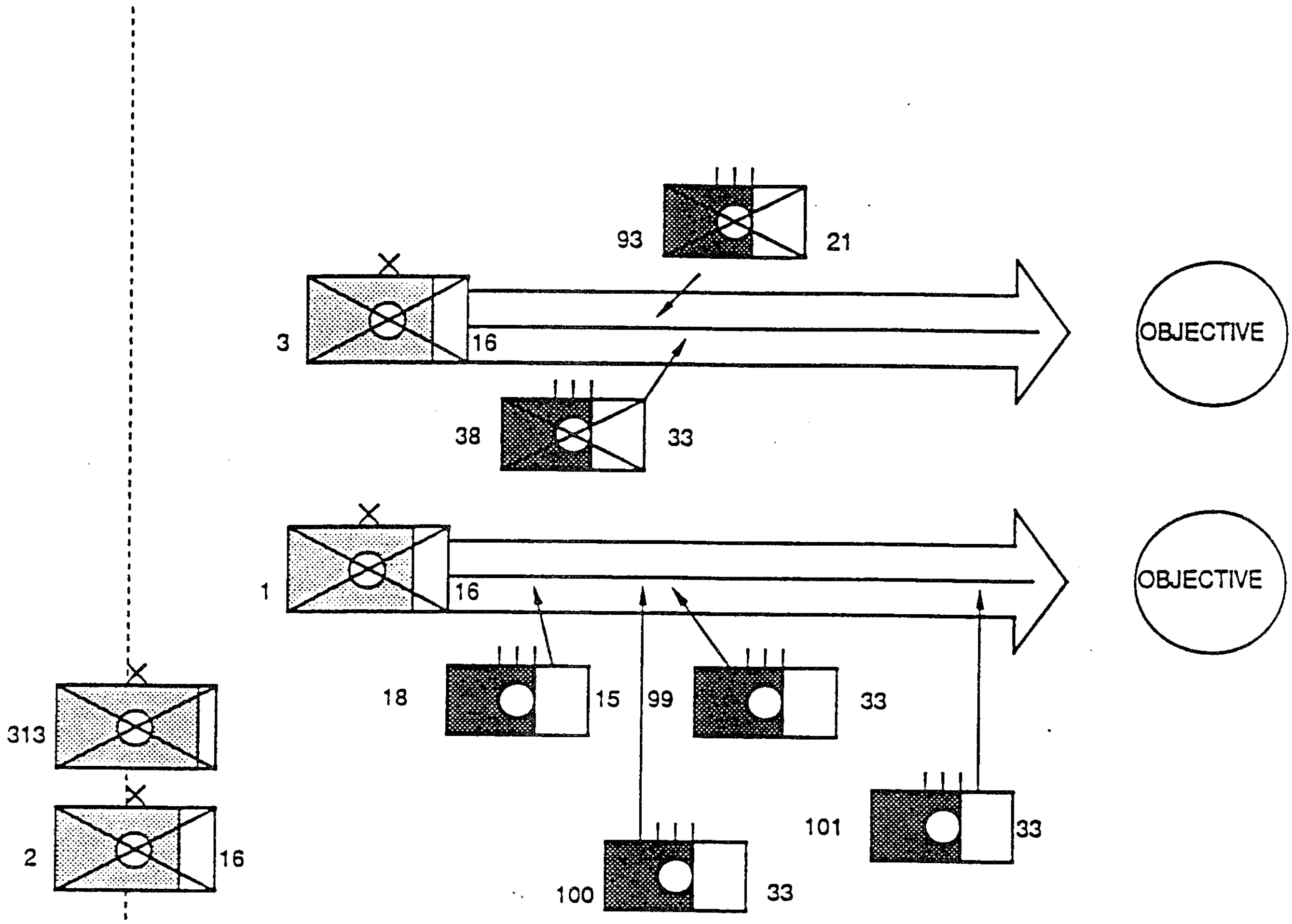


Figure 12c. Potential threats to mission: Visual method

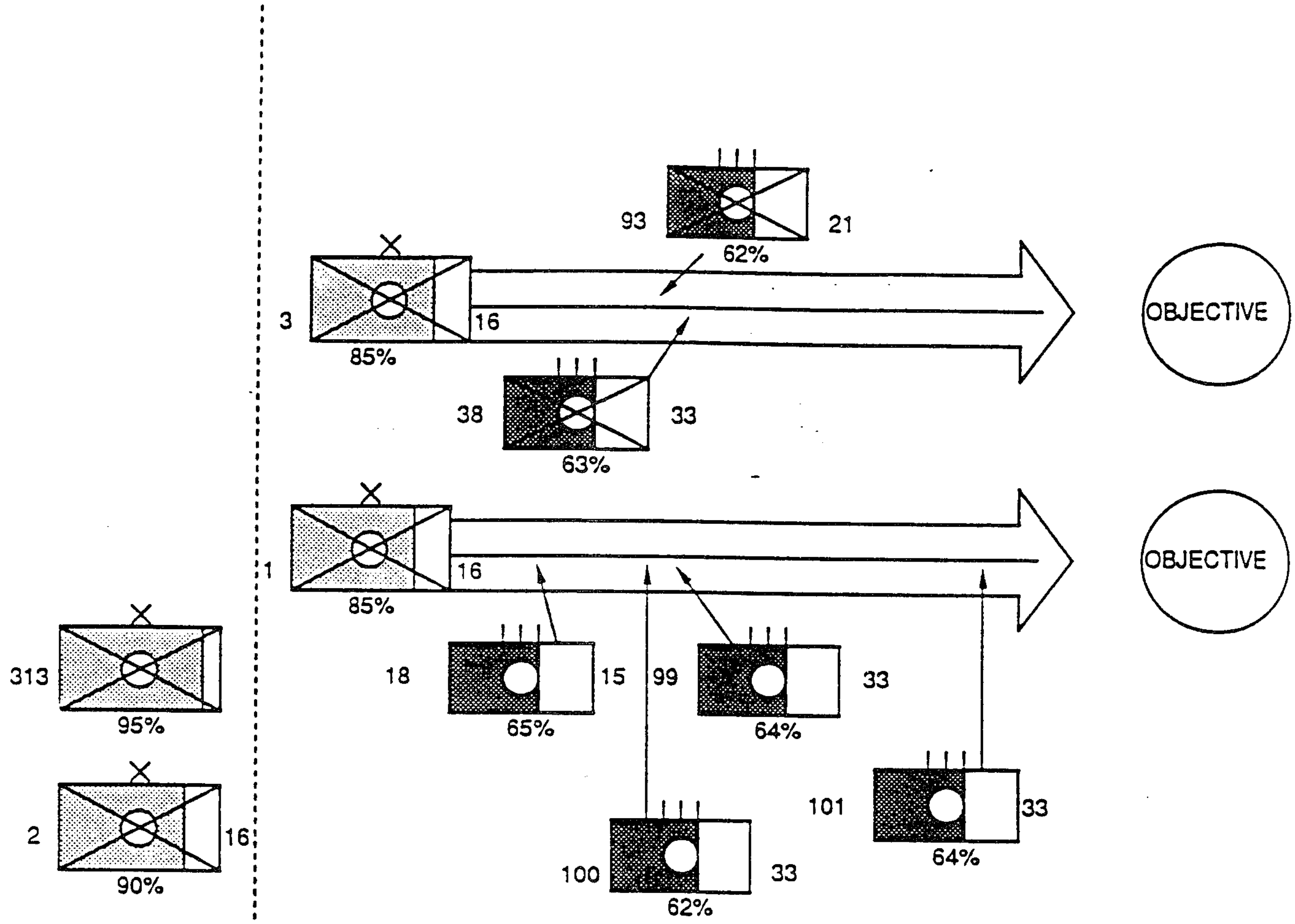


Figure 12d. Potential threats to mission: Visual plus numeric

battlefield (IPB) process. The main point of this was to illustrate how the concepts being demonstrated may be integrated into products the Army currently uses. The decision template presents enemy decision points and the potential courses of action/threats that may arise from them. Here, the Study 2 scenario was presented and showed the decision point where the 14th TA could either turn and come into the participants' sector or head south and oppose another friendly unit. In essence, this display captured the heart of the Study 2 scenario in a single graphic.

There were two variations of this display (See Figure 13 a-b). The first used standard army symbology and displayed the potential threats and decision point. The second used the visual combat power method to display unit combat power and had the strength of each unit below the symbols. This latter display presented the true size of the enemy force, which some Study 2 participants seemed to underestimate.

Ninety percent of the participants preferred the visual display over the standard display. In other words, participants preferred the visual and numerical portrayal of the threat relative to the friendly forces.

Wargaming/Course of Action Selection

A major focus of Study 3 was the integration of uncertainty management with course of action analysis. Of interest was exploring course of action analysis aiding. Participants were given abbreviated briefings of the Study 2 scenario and asked to come up with course of action recommendations.

Participants were given two variations of the "potential threats to mission" displays that depicted the friendly mission and potential threats. One variation depicted the 14th TA in sector and the other out of sector (these two conditions matched the two conditions used in the main study with the intelligence information processed). Participants were told that the enemy regiments he opposed (six in all) were each at 65%, his two lead brigades were at 85% each, one of his reserve brigades was at 90% strength and the other at 95%. When the 14th TA was in sector, participants were told it was at 95%. When the 14th TA was not in sector, participants were told to assume that it would not be a factor in their mission. The order in which these two scenarios were presented was counterbalanced across participants.

Participants made general course of action recommendations (i.e., attack with main attack north, main attack south, defend, etc.). After each recommendation, participants were then shown aiding displays appropriate to the condition they were in (which had been predetermined).

In the control condition (which corresponds to the way participants would perform in the field), no aiding was delivered. In other words, participants needed to conduct all wargaming and evaluation mentally.

The first aiding condition provided participants with probability and casualty information regarding their courses of action. Table 13 depicting

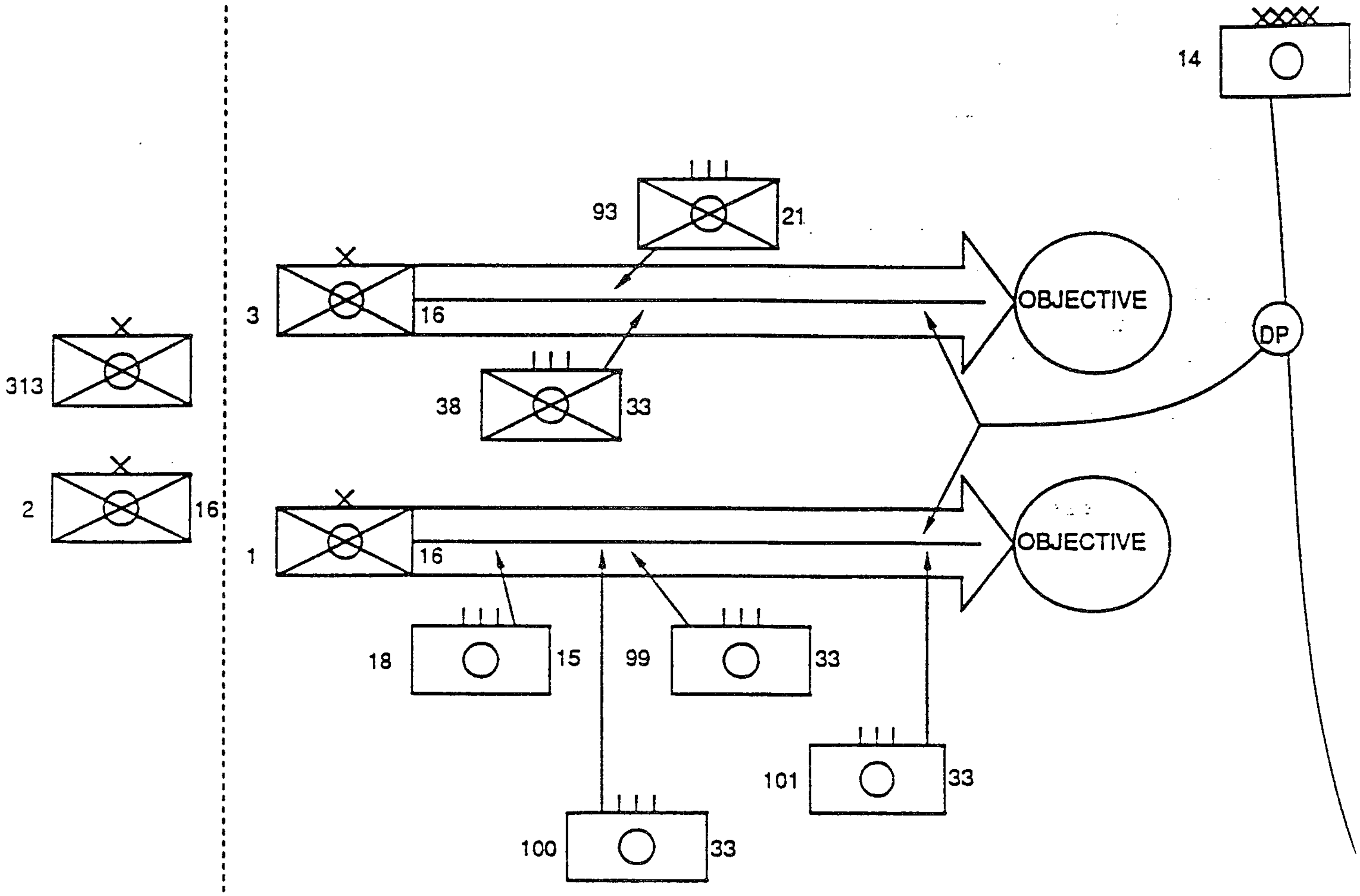


Figure 13a. Decision template: Standard symbology

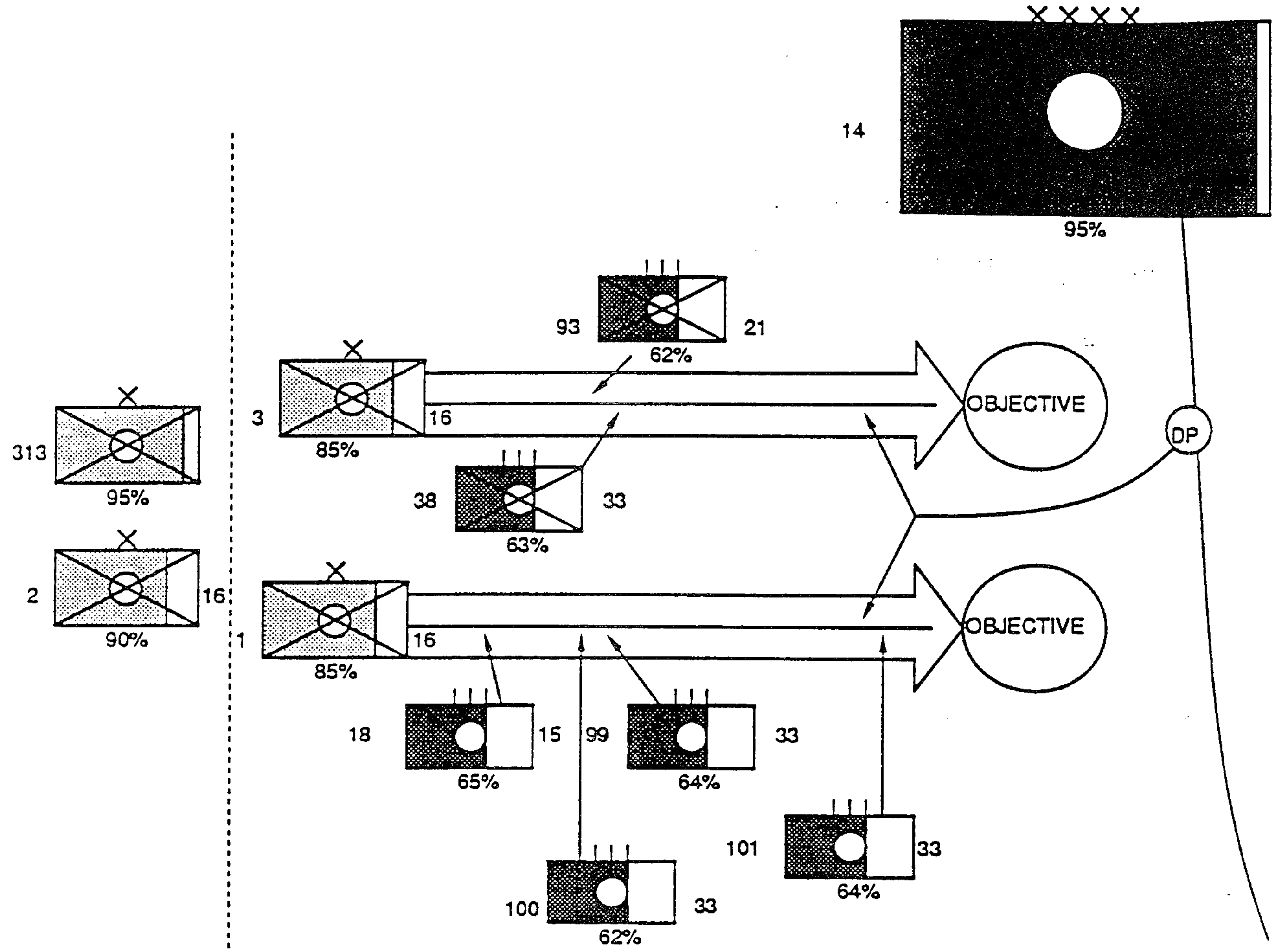


Figure 13b. Decision template with representation of combat power

probabilities of success and casualties for each course of action is presented below:

Table 13

Probability of Cuccess and Casualties for Courses of Action

	Attack North		Attack South		Defend	
	p(success)	casualties	p(success)	casualties	p(suc.)	casu.
14 TA present	.2	50%	.1	50%	.8	15%
14 TA absent	.8	25%	.5	25%	.9	10%

In defensive operations, probability of success means preventing an enemy objective, but not achieving objectives. For offensive operations, probability of success means achieving objectives. As can be seen from the table above, if the 14 TA is absent, the best COA is to attack in the north (assuming achieving objectives- -the mission--receives a higher weight than preserving the force), due to the fact that the division would have to fight through 2 regiments in the south and 4 in the north. If the 14 TA is present, the best COA is to defend given the low probability of success and the desire to preserve the force (from Study 1, participants stated that they typically desired an 80% chance of success when attacking).

The way the aiding worked here was that when participants made a recommendation, they were shown a display that was similar to the one they used in selecting their course of action, but contained the probability of success and casualty figures. Hence, participants got a projected outcome. Participants were then told that they could recommend the course of action they developed or explore another one. This process iterated until participants selected a final course of action.

The second aiding condition presented the same probability and casualty information as above, except that the aiding display also showed the visual representation of combat power and enemy strength. The intent was to enable participants to understand visually, why the numbers were the way they were.

The third aiding condition showed a "graphic" representation of the implied outcome of the course of action (See Figure 14 a-f). In other words, probability information was displayed in terms of an updated FEBA and newly positioned units. For example, when the 14 TA was present, the displays for both attacks showed the friendly units making little progress (reflecting the .1 and .2 probabilities of success for attack south and attack north). This was designed to show realistically that the division was expected to make little progress. The defend display showed the units holding in place. Similarly, when the 14 TA was not present, the defend display showed the friendly unit holding, the attack south showed the division making progress

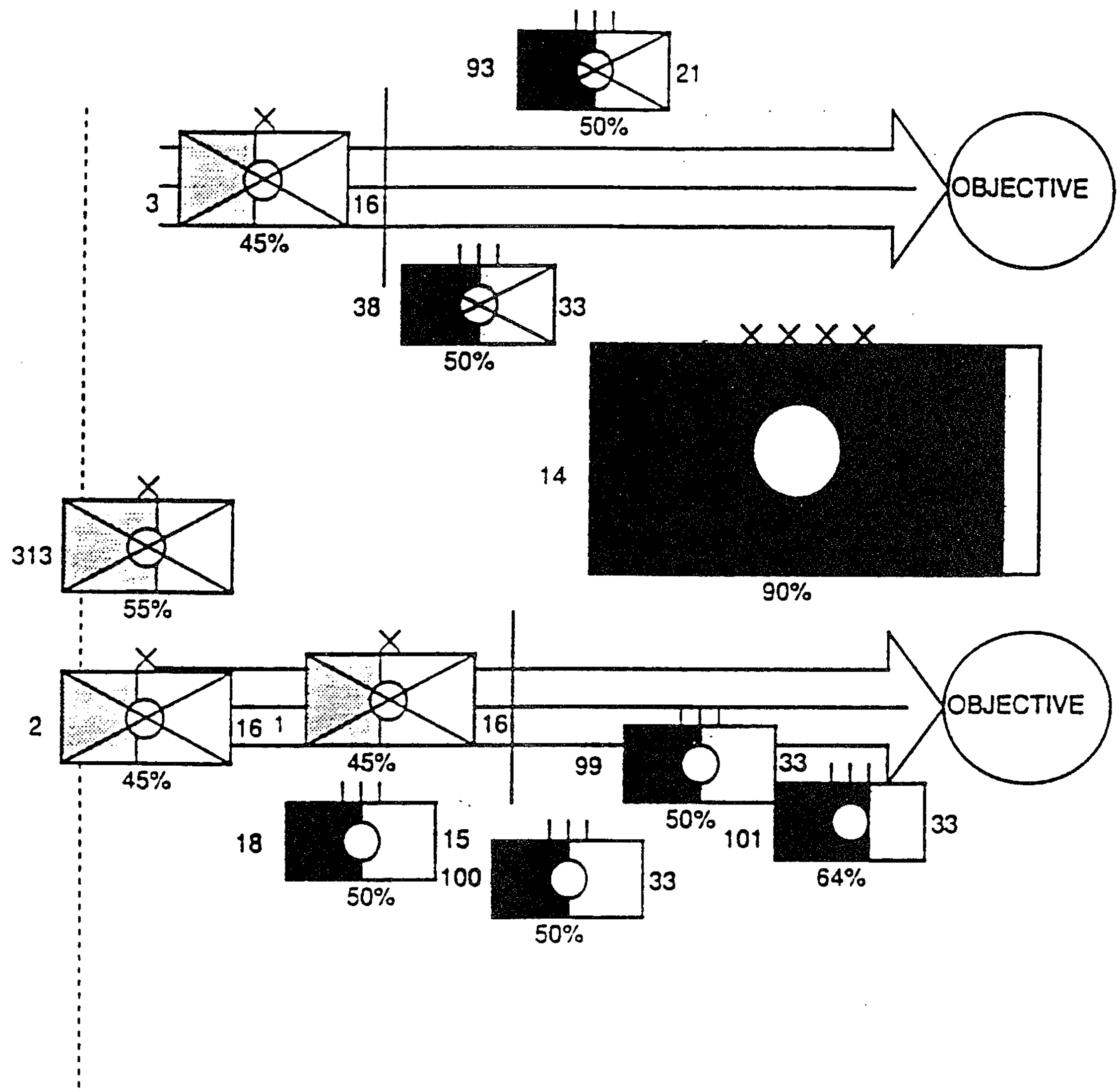


Figure 14a Wargaming: 14th tank army in sector, main attack in South.

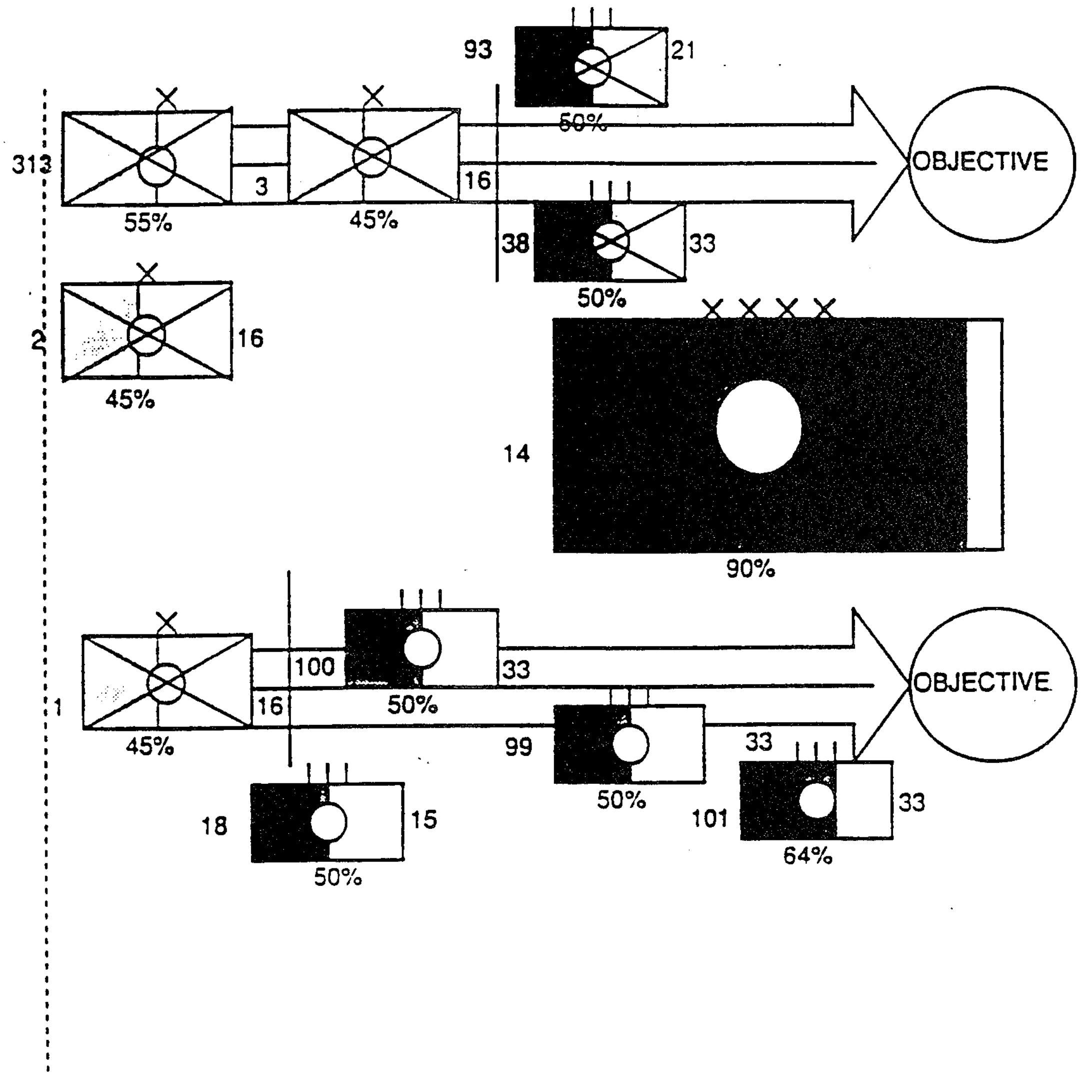


Figure 14b Wargaming: 14th tank army in sector, main attack in North.

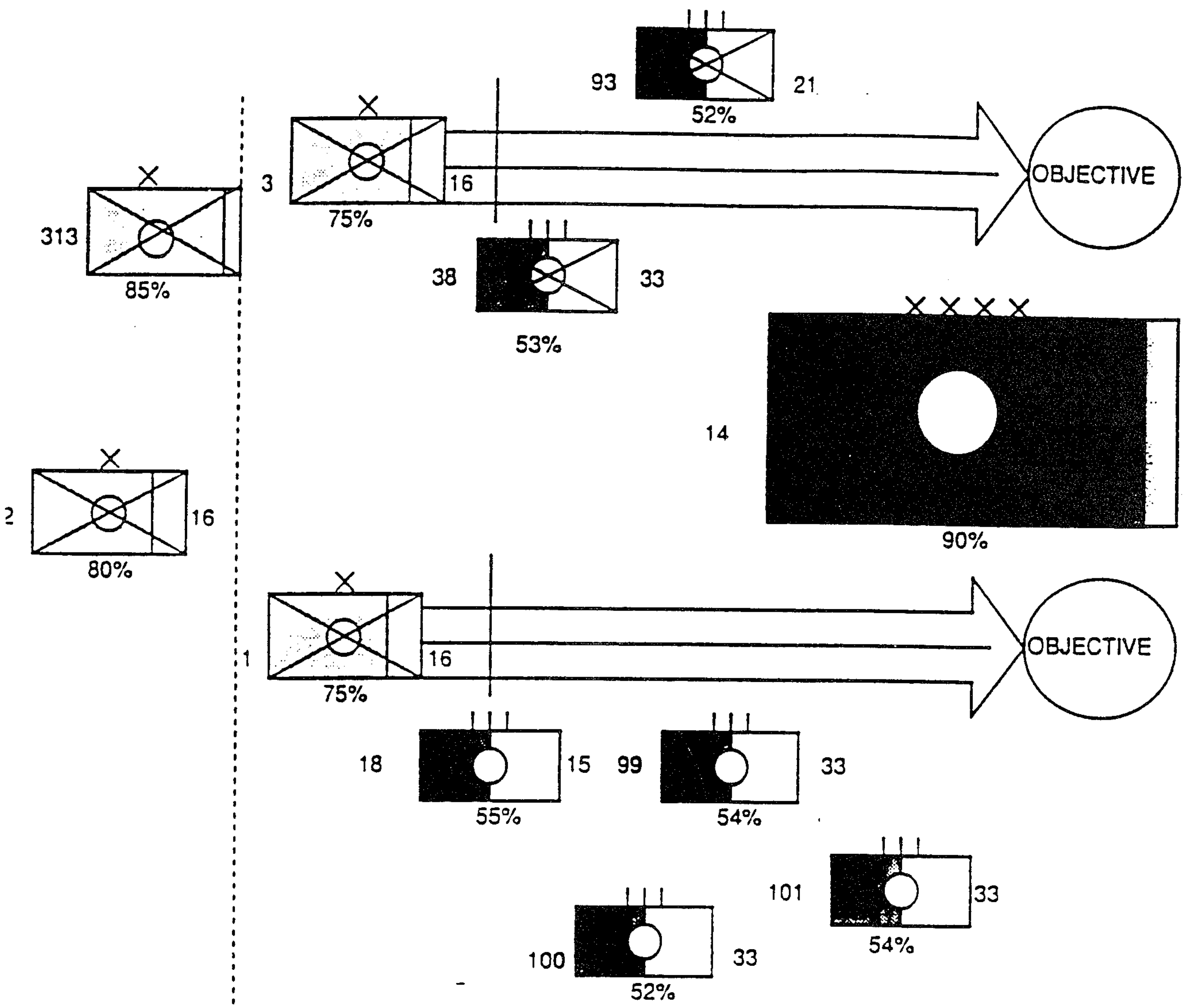


Figure 14c Wargaming: 14th tank army in sector, defend.

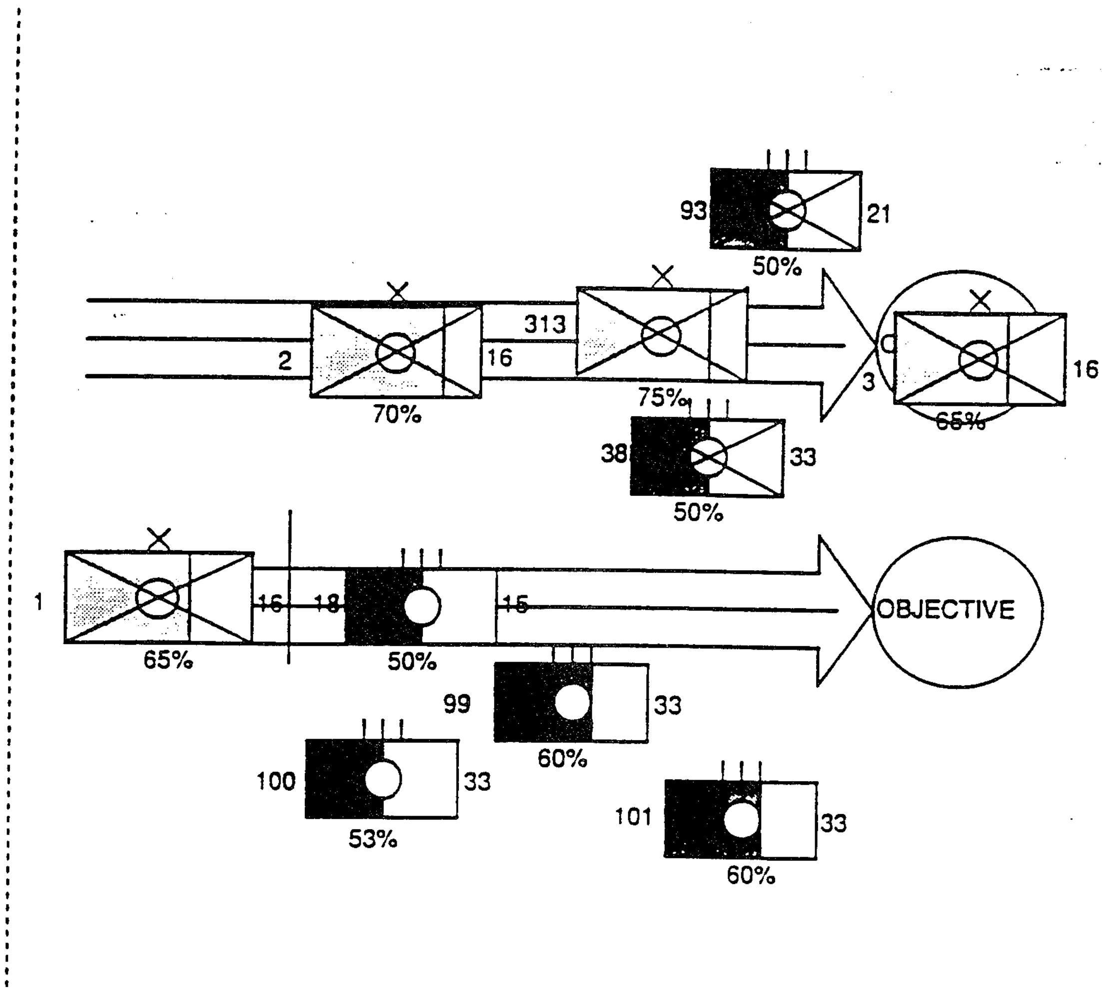


Figure 14d Wargaming: 14th tank army not in sector, main attack in North.

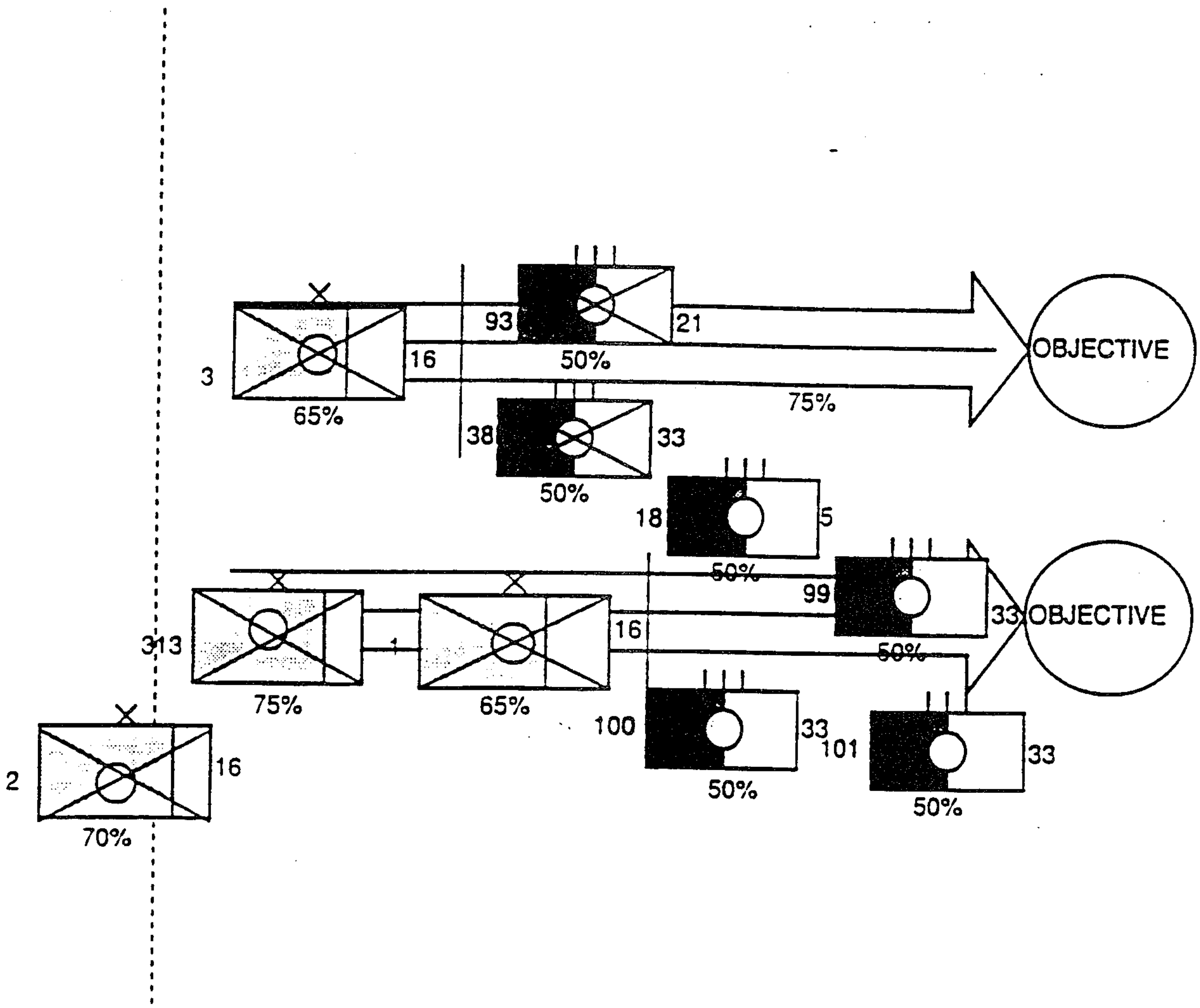


Figure 14e Wargaming: 14th tank army not in sector, main attack in South.

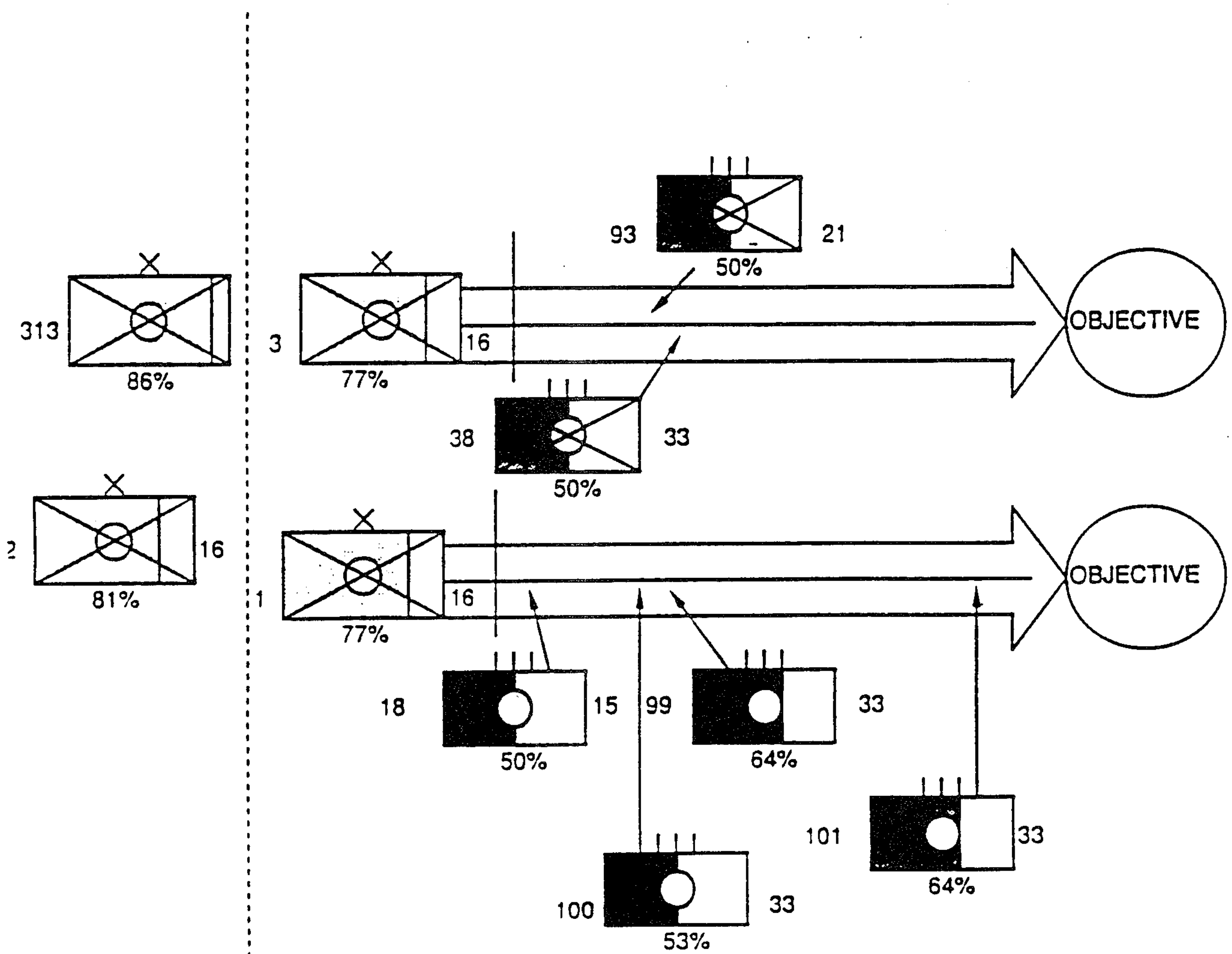


Figure 14f Wargaming: 14th tank army not in sector, defend.

but being stopped short of the objective (reflecting a 50% chance of success), and the attack north display showed the division reaching the objective in the north but being held in the south.

For this third aiding display format, casualties were factored into the combat power and strength graphics used in the second aiding condition. In other words, a unit at 85% strength that was expected to take 25% casualties, wound up at approximately 65% strength at the end of the mission.

Therefore, the second and third aiding conditions could be viewed as a "before and after" look at the mission. Both displayed combat power, but the second showed combat power and expected outcomes prior to the course of action's inception, while the third aiding concept showed the projected combat power and outcomes after conducting the course of action. Another way to dichotomize the distinction between aiding concepts two (and one) and three was that the former was a probabilistic presentation of expected outcomes, while the latter presented a likely outcome scenario.

In all aiding conditions, participants were allowed to look at as many courses of action they wanted until they picked the one they wanted to recommend.

Table 14 below presents the courses of action selected by participants under each condition.

Table 14

Courses of Action Selected by Participants

	14th TA present			14th TA absent		
	defend	attack N.	attack S.	defend	attack N.	attack S.
standard	4	1	0	1	1	3
probability	4	1	0	1	3	1
prob. + combat	4	1	0	0	5	0
FEBA + combat	5	0	0	0	5	0

(note: underlined courses of action represent the "best" choice)

The table above is interesting in that some trends appear. First, in only the FEBA update (aiding concept 3) or outcome scenario approach to aiding did participants pick the "optimal" solution every time. In the other three conditions, at least one participant chose to attack when the 14 TA was present. This 20% "attack rate" is comparable to the 33% "attack rate" of participants in Study 2 when they found the 14 TA in their sector.

When the 14 TA was not present, the two aiding concepts that depicted relative combat power produced the "best" course of action. However, even though participants were presented with probability feedback, 40% in the first aiding condition chose "suboptimal" courses of action. The participant who chose to attack in the south apparently "satisfied". Once he saw that the attack in the south had a 50% chance of success, he adopted it without further

inspection of courses of action. However, the participant that picked defend actually looked at 3 courses of action before selecting defense (his third option looked at) as his course of action. What is even more surprising is that in the control condition, only one participant picked the best means to continue the attack, and as in the first aiding condition, another participant chose to go on the defense.

The results suggest that aiding may help participants to make better decisions, particularly when the aiding helps to clarify the relative combat power of the forces involved. Presenting outcomes graphically rather than probabilistically seemed marginally preferable.

Another benefit that aiding may provide is to induce participants to consider more than one alternative course of action. Participants liked the idea of a "low cost" means to look at several courses of action and have the "aid" do all the analysis for them. In fact 80% of the participants in aiding conditions considered more than one course of action. It was hard to guess how many courses of action were considered in the control condition since participants were asked to recommend one course of action and no aiding cycles were initiated.

However, even more interesting than the finding that aiding seems to induce participants to consider alternative courses of action, was that under aiding conditions, participants often changed their minds regarding the courses of action they considered. In fact, participants adopted other than their first course of action considered approximately 37% of the time. 10 of the 15 participants in the aiding conditions changed their minds at least once in the two scenarios (14th TA present and 14th TA absent). This finding is especially interesting in light of the evidence that people tend to persevered to initial decisions, even in the face of disconfirming evidence. These findings suggest that aids can help participants to avoid this bias by showing them potential bad outcomes and allowing them to explore alternatives.

Finally, participants were asked to rank order their preferences for the different aiding concepts (after they made both recommendations, they were presented with screens from the standard and three aiding conditions and asked to provide their preferences). The results are summarized below in Table 15:

Table 15

Preferences for Different Wargaming Representations

ranking:	1st	2nd	3rd	4th	mean
standard	0	0	3	14	3.82
probability	1	2	13	1	2.82
prob. + combat	10	7	0	0	1.41
FEBA + combat	8	8	1	2	1.84

As can be seen from Table 15, there was overwhelming preference for some type of aiding (no participant ranked the standard format as their preference). There was strong support for including aiding that depicted combat

power (using the visual format). Participants were generally split as to whether they wanted outcome information probabilistically or in scenario outcome format (FEBA updating), although there was a slight preference for probabilistic representations. Again, two participants explicitly volunteered that they would like to see both.

The results of the course of action analysis aiding suggest some important things: first, that aiding in this area is desired by participants. Second, to the extent that our conclusions about decision quality are correct, aiding appears to improve decision making by providing outcome feedback regarding considered courses of action. Third, aiding seems to induce participants to consider alternatives and even change their minds if necessary. As such, aiding may improve decision making and reduce biases.

DISCUSSION

Aiding appears to be desired by participants and needed. Participants are particularly interested in intelligence database aids and wargaming aids, especially, those that integrate the two. Participants typically had a preference for displays that conveyed the most information, especially if it was understood that they would not have to provide the information that goes into the aid (i.e., the participants liked automated aids where they don't have to do the work to build the database).

Participants were most "purist" about the maps and overlays they used, wanting little change there. However, they were open to new methods of displaying information in automated or other media. This is consistent with traditional resistance to major organizational changes. Aids that "teach" existing procedures are likely to be resisted least.

Aids appeared to have a "stimulating" effect in that they induced participants to think more about the problems they were working on and even consider and adopt new courses of action. This alone may make aiding worthwhile, i.e., encouraging breadth of thinking as well as supporting depth, thus, of course, tradeoffs with the urgency of time stress.

Finally, participants differed in their preferences, both across each other and across contexts. Therefore, aids seem to be required that present information in a variety of formats and enable participants to pick and choose what they want to see in a way they want to see it (with perhaps some alerts to protect participants from using aiding to support biased thinking). Also, aids need to be sensitive to the fact that different cells in the decision process (both laterally and hierarchically) have different requirements and aids need to be sensitive to these.

STUDY 4

INTRODUCTION

The results of Study 3 suggest that the approaches to aiding tested show promise, both in terms of enhancing decision making performance and user acceptance. This is particularly true of the two aiding concepts identified as principal needs in Study 2.

Study 3 suggests that aiding the management of intelligence information process through interactive displays is highly promising. The data showed that such interactive displays were well received by Study 3 participants. In addition, Study 3 results suggested that graphic and numerical displays of such information (both in terms of unit strength and relative combat power analysis) actually improve the accuracy of judgments regarding this information.

This improved judgment appears to have important implications for decision making behavior as well. Study 3 data show that when this intelligence information is combined with outcome information (produced from a "wargaming aid"), participants make more appropriate decisions given the tactical situation. Together, the intelligence aiding and wargaming aiding improve both judgment and decision making in the context of the Study 3 pilot. An additional by-product of this aiding approach was that participants considered more alternative courses of action before making a final recommendation, and often were willing to pick other than their first course of action considered--something unaided participants were less inclined to do.

Aiding

The purpose of the present study is to investigate the effects of aiding in a more realistic and richer scenario than the examples used in Study 3. In addition to the "computerized" aiding format (which the Study 3 aids were designed to simulate), another aiding approach is investigated. Throughout the present series of studies and in previous ones (cf. Leddo et al., 1987), study participants have requested additional people to help them in their decision making and to assume the roles of other staff members with whom they would normally work.

Groups versus Individuals

The issue of group versus individual decision making is an important one in this series of studies. From an experimental point of view it can be argued that the previous studies did not hit the organizational issues earlier discussed in Study 1 in that they focused on individual decision making--something that typically does not occur alone in the tactical planning context. Therefore, it could be argued that the organizational decision making process could not be fully understood, nor the effects of aiding, unless decision making is examined in the organizational context.

The present study begins to address this issue. First, it examines whether aiding can help decision makers under uncertainty and time stress. Second, it examines whether alternative approaches to manpower-based (i.e.,

increase the number of people working on the problem) aiding can be as effective as adding additional people to the decision making process.

The present study employs a 2 x 2 design. The first factor pertains to whether decision makers receive aids other than additional people. Half of the participants in this study receive no such aiding. The other half receive two types of aids. The first is an intelligence information management aid that organizes and displays situational information about the enemy. The second is a wargaming aid that shows the outcomes of different courses of action in terms of FEBA (Forward Edge of the Battle Area) movement and casualties.

The second factor pertains to whether decision makers receive additional people to assist in the decision making effort. Half of the participant cells in this study work on the problem alone (as in Study 2) and the other half of the cells are comprised of pairs. Since the G-2 (intelligence officer) is the officer that most closely works with the G-3 (the focus of this series of studies) and is the officer most requested by participants in the previous studies, the present group decision making cells will be represented as a G-3/G-2 pair to represent the more realistic organizational context.

Hypotheses to be tested

There are several hypotheses to be tested in the present study. The most important hypothesis relates to whether aiding enhances decision making. There are two ways in which decision making can be enhanced: it can be made faster or lead to better decisions. Ideally, decision aiding would have both effects. The worst case would be for aiding to lead to both slower and poorer decisions. If there was a split, i.e., decisions were slower, but better or faster but poorer, it would require a cost benefit analysis in order to examine the tradeoffs. (It can be argued that currently decisions are of sufficient quality. If so, the time stress variable would reduce quality below that required by reducing the capability to fully address the information as would be done in non-stressed conditions. If this is the case, speeding the process from a degraded decision to the current quality would be preferred to increasing quality from its non-time stressed level to a higher level. This argument is too complicated to pursue here.) For the purpose of the present study, the hypothesis is that both the aiding condition and the group condition will show main effects in improving decision making in terms of both speed and quality of decisions.

Of particular interest is comparing the group (The term group will be used here to represent the G-2/G-3 organizational pair.) unaided condition with the individual aided condition. This condition has important implications as to whether aiding via non-human means can produce organizational performance comparable to adding additional personnel. In the present study, it is hypothesized that the individual aided condition will lead to comparable performance (i.e., no significant difference) to the group unaided condition both in terms of speed and quality of decision.

Finally, Study 3 found that use of a wargaming aid led to participants considering more courses of action than did participants who had no aid. It is hypothesized that participants in the aided conditions will consider more courses of action than those in the unaided conditions.

METHOD

Participants

There were 56 participants in Study 4. Participants ranged in rank from Chief Warrant Officer 2 (CW2) to lieutenant colonel (LTC). All participants were active duty officers from either Ft. Carson, Ft. Polk or Ft. Hood. Participants were from a variety of branches including combat arms, military intelligence (MI) and combat support.

Materials

The scenario used in Study 2 was used in Study 4. The version used was the "Good to Bad" one where participants were initially instructed that the 14th TA was expected to commit itself in a different sector but winds up committing itself in the participants' sector. The rationale behind selecting this version was that participants' course of action selections in Study 2 were strongly influenced by whether they located the 14 TA. Recall that when the 14 TA was located, two-thirds of the participants chose to defend, whereas when the 14 TA was not located, 80% chose to attack. On the other hand, in the "Bad to Good" condition, whether the participants located the 14 TA had no significant effect on course of action chosen. It was decided that it would be desirable to use a scenario where course of action selection had been shown to be tied to how well participants resolved uncertainties.

Intelligence Aid

Displays were constructed to simulate an interactive intelligence information management aid. The intent of the aid was to allow the user to call up information in a way that was most useful to him. In addition to the type of information, the aid allowed the user the flexibility to call up information by unit, sector or across the battlefield so that he could make quick comparisons and quick assessments. The "aid" presented four types of information on the enemy that was current as of 060100 September (i.e., was five hours old from the scenario start time). These types included order of battle information, previous activity, similar activity, and enemy intention.

The order of battle information presented enemy unit identifier, its type and echelon (e.g., tank regiment), its parent unit, its strength (in both visual and numeric format from Study 3), its disposition, and its current activity. Participants could call up this information for any level of aggregation or echelon (e.g., by division, regiment).

The previous activity displays depicted what the unit presented was doing 12 hours prior to the latest update. This enabled participants to compare the previous activity to the current activity in order to project the "flow of the battle." Participants could call up this information by different levels of aggregation or echelon.

The similar activity displays allowed the user to compare what units were engaged in common activities such as attacking, defending, resupplying, etc. This information could be called up by unit, sector or across the FEBA. The purpose of this information was to allow the user to get a sense of how the enemy was coordinating his forces.

The enemy intention displays presented the G-2's estimate of the enemy's most likely course of action over the next 24 hours. For participants in the group conditions, they were told that the estimate was derived from the previous shift (at 060100 September). This was done because the group conditions had a participant role playing the G-2. The enemy intention could also be displayed by units at different levels of aggregation and echelon. The intelligence aids used in Study 4 are similar to those presented in Study 3 (See Figure 11a-d).

Wargaming Aid

The wargaming aid used in the present study was a scaled-down version of the FACT (Force Allocation and Comparison Tool) wargaming aid (Vane, Black, Laskey, and Bresnick, 1990). It was programmed in Hypercard on an Macintosh desktop computer that was carried to the sites where the study was conducted.

The wargaming aid embodied the tactical situation as of 060100 September. This included the initial friendly and enemy strengths, locations, and activity as well as the two axes of advance and the objectives.

The user could take this initial information and update enemy strengths, locations, and missions based on any intelligence information that had come in since the initial situation, including assigning elements of the 14 TA. In addition, the user could update friendly strengths, locations, and missions based on new information and courses of action being considered. Assignment of missions done for each of the two axes of advance, thereby giving the user the flexibility to attack along one axis and defend along the other, assign a main and supporting attack, etc.

Once the user updated the tactical information, he could run the wargame to obtain a projected outcome of the battle. The outcome information was presented both in terms of casualties (expressed in unit strengths resulting after the battle) and updated FEBA movement (i.e., which side made progress). The user could then see whether his defense would hold or whether his attack would seize the objectives based on the movement (or lack thereof) of the FEBA. The aid was constructed so that the user could see the results of as many courses of action as he wished, thus enabling him to survey several options. The displays presented in the aids are similar to those presented in Study 3 (See Figure 14a-f).

The calculations the aid used in order to arrive at FEBA movement and casualties were based on a force on force analysis derived from formulas discussed in the Army doctrinal publication ST100-9.

Procedure

Participants received a background questionnaire upon arriving at the study site. The questionnaire contained questions relating to the participants' branch, rank, experience, and training. The purpose of the questionnaire was to determine the suitability of participants to role play a G-3 or G-2. In cases, where two officers were used, the questionnaire (plus follow-up discussions with the participants) served as the basis of deciding which role each participant would play. The decision whether to run the individual or group condition was driven by whether the host installation had scheduled one or two participants. Participants were assigned to the aided (i.e., receiving the intelligence and wargaming aids) or unaided conditions on a random basis, with the provision that equal numbers of aided and unaided cells be run.

The fifty six participants were assigned nine per cell to the two group conditions and the individual/aided condition. Eleven participants were originally assigned to the individual unaided condition. However, two of the participants in the latter condition indicated that they had to rush through the problem to meet other commitments. Therefore, the data of these two participants is not included in the analysis, leaving nine observations for each of the four cells conducted.

The procedure and data coding process was similar to that used in Study 2. There were three major differences. First, the sessions were not videotaped. Second, participants in the aiding conditions were given instructions on the aids that were available. Rather than have the participants operate the aids themselves, they were instructed to tell the experimenter what they wanted and the experimenter operated the aid (or turned to the appropriate display in the case of the intelligence aid).

Because aids were introduced into the present study, the experimenter also recorded when and how much time participants used the aid. This is an important source of data in its own right. A big question in the aiding field is whether decision makers will use the aids that are developed for them. Therefore, simply because a participant has an aid made available does not mean that he'll use that aid.

A final measure was also taken in the present study. Since one of the hypotheses is that participants in the aiding conditions will consider more courses of action than those in the unaided conditions, a measure was taken on the number of courses of action considered. This measure was a composite of the courses of action discussed as considered in the final briefing, the courses of action discussed by participants in their think aloud during problem solving, and the courses of action wargamed using the aid.

The third major difference in the procedure between the present study and Study 2 is the elimination of the time stress condition. One of the major hypotheses to be tested in Study 4 is that aiding and group decision making leads to faster decisions than unaided and individual decision making. Therefore, rather than giving participants a fixed time limit (as was done in the time stress condition of Study 2), participants were allowed to work at their own pace. The dependent measure regarding time was the amount of elapsed time (recorded by stop watch) between the conclusion of the initial

scenario briefing and the point at which participants announced they were ready to make their final briefing.

RESULTS

Treatment of Participants Use of the Aids

Not all participants used the aids provided. In the individual condition, only five of nine participants used the aids. The primary reason given by those who did not use the aids was that they were unfamiliar with them. As such, this percentage may not be reflective of those who would not use aids under battlefield conditions if they had extensive experience with this aids in peace time. Those that did use the aids said that they accepted the information and conclusions presented to them and assumed they were valid. This in itself is an interesting finding suggesting that developers of aids need to be careful to validate their combat calculi, lest decision makers blindly accept the output of such aids as gospel. However, it is unclear whether such decision makers would readily accept the output of the aids if there were lives on the line as opposed to a constructed experimental setting with no real-world consequences to their decisions.

In the group conditions, eight of nine cells used the aids. This higher percentage is consistent with the individual data. If 44% of individuals are reluctant to use an aid, then assuming independence, only 19% of the time would a pair of individuals be comprised of two people who are both reluctant to use aids. This is consistent with the 11% of groups in the present study (given the sample size of the present study) that contained two members who did not use the aids. An additional factor that may have reduced the number of groups that did not use aids may be that in groups, neither individual wanted to appear to the other that he did not feel comfortable with a new aid.

The finding that groups showed more overall aid usage than individuals has practical implications for the introduction of aids into the decision making process. Specifically, aids that help group decision making are more likely to find proponents within the decision process than aids targeted for a single user. Even if only one person in a group uses the aid, the aid still has the opportunity to impact the group decision.

Given that five of the 36 cells in the present study did not use the aids, the question remains as to what to do with their data. Since these five cells were, in fact, "unaided," their data is not appropriate in the aiding conditions where the issue is how aiding affects decision making. However, it does not seem appropriate to include the data in these cells with the unaided conditioned given that the former do not represent a random sample of unaided people. Rather they are a sample of people who had the opportunity for aiding and turned it down. Therefore, the comparison of aided versus unaided conditions will be analyzed without using the data of these participants. However, for most comparisons (as appropriate), a separate analysis that includes the aided-not used participants as a separate condition will be done both for comparative purposes and for completeness.

Times Required to Arrive at COA Recommendation

The first major hypothesis tested in this study is that both aided and group conditions would show a main effect in decreasing the amount of time to arrive at a course of action recommendation.

Table 16 below shows the mean decision times in minutes for participants in each of the four experimental cells.

Table 16

Time taken (minutes) to arrive at COA recommendation

	Aided	Unaided
Group	114.0	112.5
Individual	108.5	92.2

Table 17

Time taken (minutes) to arrive at COA recommendation

	Aided	Unaided	Aided, Not Used
Group	114.0 (N = 8)	112.5 (N = 9)	155.3 (N = 1)
Individual	108.5 (N = 5)	92.2 (N = 9)	81.6 (N = 4)

Overall Mean 104.93

Source	DF	F	P
Group-Individual	1	4.45	.04
Aided-Unaided-Not used	2	0.62	.55
Interaction	2	1.25	.30

In Table 16, the differences in means are not statistically significant. (All F's for main effects and intentions non-significant). Table 17 displays the same results with the added 2 cells for the aided-not used condition. Included in each cell is the mean and in parentheses the number of observations. Also included in Table 17 are the sources of effects with associated degrees of freedom, F scores, and probabilities. As indicated, the main effect for Group versus Individuals is significant at the .05 level (P = .04) with the aided-not used cells added. Further examination of the table indicates that the single group that refused to use the aid had a time well above the overall average to reach a course of action decision. Further the mean time for the individual-aided-not used cells was well below the overall mean. As earlier indicated, this finding is difficult to interpret because of

the unique nature of the participants who refused to use aids. Clearly, these participants are not a random sample of participants. A stereotypical view of individuals who are "averse to using computers" could lead to speculation that they are likely less enthusiastic about the whole task. (Informal review of the background information on these participants revealed no obvious common characteristics or differences from those who did use the aids given the chance.) Thus, a qualified conclusion is that individuals perform the task faster than groups, but the qualification is quite important here. The clear results are that aids do not speed up the decision process, nor does using a two person combination as opposed to an individual.

As noted earlier, of particular interest to the present study is investigating whether non-human aids (e.g., computer-based) can substitute for human aids (i.e., additional staff). Here, the individual aided participants were at least as fast as the group unaided participants suggesting that individuals with machines may be as fast as groups without them. Considering that the group uses twice the manpower of the individual, the individual aided condition required half the number of person-minutes to arrive at a decision even though the elapsed time was equivalent.

Quality of Decisions Made

The second major hypothesis tested in Study 4 is that aided and group decision making leads to better quality decision making. Study 2 demonstrated that the course of action selected by participants was influenced by whether or not they had located the 14 TA (resolved that major part of situational uncertainty). Given that one function of aiding or using additional people might be to help resolve uncertainty, it was decided to look at whether participants resolved uncertainty (located the 14 TA) as a function of whether they received aiding or were in a group setting.

Table 18 presents a breakdown of ability to locate the 14th TA as a function of group versus individual conditions (excluding "not used"). The result is statistically significant at the .05 level ($\chi^2 = 5.10$, df = 1, $P < .05$) Groups perform better than individuals.

Table 18

Ability to Locate the 14th TA as a Function of Group versus Individual

	Locate 14th TA	Did Not Locate 14th TA
Group	11	6
Individual	4	10

Table 19 displays the results with the aided-not used participants included. The results are likewise significant at the .05 level ($\chi^2 = 5.5$, df = 1, $P < .02$). Again, the result is confirmed.

Table 19

Ability to Locate the 14th TA as a Function of Group versus Individual with "Aided-not used" Participants Added

	Locate 14th TA	Did Not Locate 14th TA
Group	12	6
Individual	5	13

Table 20 shows a breakdown of the ability to locate the 14th TA as a function of the aided versus unaided conditions with "aided-not used" deleted. The result is statistically significant. Participants in the aided condition performed better ($\chi^2 = 3.89$, df = 1, $P < .05$). Table 21 shows the analysis with the "not used" participants added as separate cells. This result has a Fisher's Exact test probability of .05, which is statistically significant. Table 22 shows a third view with the "not used" added to the unaided cell. This result is also statistically significant ($\chi^2 = 3.95$, df = 1, $P < .05$). Thus, all three views agree on the impact of aiding on the ability to locate the 14th TA.

Table 20

Ability to Locate 14th TA as a Function of Aided versus Unaided

	Locate 14th TA	Did Not Locate 14th TA
Aided	9	4
Unaided	6	12

Table 21

Ability to Locate the 14th TA as a Function of Aided versus Unaided with the "Aided-not used" Cells Included as Separate Cells.

	Locate 14th TA	Did Not Locate 14th TA
Aided	9	4
Unaided	4	12
Aided No Used	2	3

Table 22

Ability to Locate the 14th TA as a Function of Aided versus Unaided with the "Aided-not used" Included

	Locate 14th TA	Did Not Locate 14th TA
Aided	9	4
Unaided	8	15

Table 23

Number of Cell Entries Locating or Not Locating the 14th Tank Army

	Aided		Unaided	
	L14TA	~L14TA	L14TA	~L14TA
Group	6	2	5	4
Individual	3	2	1	8

Returning to Table 20, while the data show main effects for aiding and number of decision makers, these effects appear to be the by-product of an interaction. As displayed in Table 23, the majority of observations in each condition located the 14 TA, except the individual/unaided condition where only one of nine participants located the 14 TA. Using Fisher's Exact test, the results for the group/aided conditions and the individual/aided are respectively .327 and .066, the latter near statistical significance at the .05 level.

These data provide support for the contention that providing aids or additional human support can lead to improved uncertainty resolution without increasing time required. The individual/aided participants did about as well in locating the 14th TA as the group aided, suggesting that in the context of the present study, non-human-based aiding may be just as effective as human-based aiding.

Number of Options Considered

A secondary hypothesis of the present study is that aiding leads to participants considering alternative courses of action. Study 3 suggested that unaided participants will often go with their first course of action, even when other options may be better, whereas aided participants will often consider multiple options before picking the best one.

On average, participants in aided conditions considered 4.3 courses of action, whereas those in unaided conditions considered 1.8 courses of action. This difference was significant, $F(1,27) = 10.17$, $p = .004$, thus supporting the hypothesis. (See Table 24 for results; "aided-not used" participants are not included here.)

Table 24

Number of Options Considered by Condition

COA's considered

Aided	4.3
Unaided	1.8
Group	3.3
Individual	2.4

While this was not part of the original set of hypotheses, the number of options considered by groups and individuals was also examined. On average groups considered 3.3 courses of action, while individuals considered 2.4 courses of action. This effect was not statistically significant.

Courses of Action Chosen

The main variable of interest in the present study was the courses of action chosen by participants as a function of whether they participated individually or in groups and whether or not they used aids. Given that in Study 1 it was found that whether participants located the 14th TA (resolved uncertainty) had a large effect on courses of action chosen, this variable was also included in the analyses below.

In order to evaluate the quality of course of action recommendations, it must first be determined what courses of action are appropriate in the present scenario. Given that the scenario used was novel and had no Army-approved solution, it is difficult to make precise determinations as to how good individual courses of action are. However, given the specifics of the situation, some general conclusions can be reached.

The present scenario depicted a U.S. mechanized infantry division with four brigades and an approximate strength of 86% going up against three enemy divisions and a full strength tank army. Of the three enemy divisions, two were admittedly reduced to single regiments of about 60% strength and hence were largely combat ineffective, while the 33rd GTD was comprised of four regiments that were about 70% strength. The 14th TA was comprised of three divisions, each at about 95% strength.

In other words, a case could be made that the enemy force was substantially stronger than the friendly force. U.S. Army doctrine commonly calls for the friendly forces to have a 3:1 force advantage in order to launch an attack. Here, the enemy was closer to having the 3:1 advantage--a situation that Army doctrine normally calls for defense. Therefore, it is argued that in the present scenario, given the relative strengths of the friendly and enemy forces, the general course of action most appropriate to the situation is defend. In particular, given that the scenario outlined two avenues of approach into the friendly sector, it is argued that an appropriate defense should at least defend both avenues.

Therefore, for the analyses presented below, courses of action were categorized in terms of whether they defended both avenues of approach or did not defend them. Within these categories, there were numerous variations. For example, some participants chose to defend both axes and set up another defense stronghold to prevent the enemy from penetrating through a different location. These participants were, in effect, defending three "avenues of approach" which was considered in the same spirit of defending two avenues of approach.

For participants in the "other" category, some chose all out attacks (along both avenues of approach), some chose to attack on one avenue, and still others developed a hybrid approach of defending along one axis and attacking along the other.

Table 25 shows the number of participant cells that chose a course of action that defended both avenues of approach versus those that did not (aided-not used deleted). The courses of action are broken down by whether participants were tested individually or in groups, whether or not they were in the aided versus unaided condition, and whether or not they located the 14th TA. Recall that if the 14th TA was not located, the initial briefing indicated that it would not be in sector, and thus the friendly forces faced the 33 GTO at about 70% strength, roughly even to the friendly force.

Table 25

Course of Action Recommendations (Expressed as Number Choosing to Defend/Total)

	Aided		Unaided	
	L14TA	-L14TA	L14TA	-L14TA
Group	0/6	0/2	4/5	3/4
Individual	2/3	0/2	1/1	1/8

As can be seen from Table 25, the results of the individual/unaided condition replicated the Study 2 results. Specifically, the one participant who located the 14th TA chose to defend which is comparable to the 67% of participants in the Study 2 Good-Bad condition who chose to defend when they located the 14th TA. Similarly, seven of eight participants who did not locate the 14th TA did not defend (and therefore chose to attack in some form or another), which is comparable to the roughly 80% of participants in Study 2 who chose to attack when they did not locate the 14th TA. Hence, these results support the Study 2 conclusion that in the face of unresolved uncertainty, individual participants act as though they are making best case assumptions regarding that uncertainty.

The results of the individual/aided condition were similar to those of the individual/unaided condition. Specifically, of the three people who located the 14th TA, two of them chose to defend (67%) which is comparable to

the unaided condition in the present study and also to the Study 2 results. Both participants who did not locate the 14th TA chose to attack, again comparable to the present study individual/unaided and the Study 2 results. This result is not surprising since when participants run the wargaming aid without plugging in the 14th TA, it is possible to generate successful offensive courses of action.

Given that the results of the single decision maker condition cells are very similar, the hypothesis that aiding leads to better decision making is not supported. However, as noted above, participants in the aided condition were more likely to locate the 14th TA than were those in the unaided condition, suggesting that the benefit from aiding comes in increased likelihood of resolving uncertainty and developing a more correct picture of the enemy situation.

For purposes of completeness, Table 26 presents the course of action as a function of aided, unaided, or aided-not used. Fisher's Exact test probability for this result is .07, nearly statistically significant. However, as can be seen, the major difference in the table with respect to COA is for aided, where many more participants chose to attack.

Table 26

Cource of Action as a Function of Aided versus Non-Aided

	Attack	Defend
Aided	11	2
Unaided	9	9
Aided, Not Used	2	3

Table 27 displays the COA's as a function of Group versus Individual with aided-not used participants included in appropriate cells. This result is clearly non-significant. (χ^2 = .47, df = 1, P < .50).

Table 27

Course of Action as a Function of Group versus Individual

	Attack	Defend
Group	10	8
Individual	12	6

Again, for purposes of completeness, a comparison of course of action versus success in locating 14th Tank Army is shown for all participants in Table 28. These overall results are not statistically significant (χ^2 = 2.68, df = 1, P = .10). For the "aided-not used" subjects, those who located the 14th TA chose to defend. Two of these who did not locate the 14th TA chose to attack.

Table 28

Cource of Action as a Function of Locating the 14th TA (all participants)

	Attack	Defend
Locate 14th TA	8	9
Did Locate 14th TA	14	5

Given that the results of the individual/aided and unaided conditions are very similar, these cells are collapsed for the remainder of the analyses. While aiding did not influence what decisions were made by individual participants, as in Study 2, whether participants located the 14th TA did affect course of action selection. Specifically, of the four participants who located the 14th TA, three chose to defend and one chose to attack. Of the ten participants who did not located the 14th TA, only one chose to defend both avenues of approach. This difference was statistically significant, (Fisher's Exact P = .041). This suggests that at the individual level, uncertainty resolution, not aiding per se, affects course of action selection (although aiding influences uncertainty resolution).

The results of the group/unaided condition did not mirror those of the individual decision maker conditions. While it is true that four of the five participants who located the 14th TA did chose to defend (comparable to the individual conditions and Study 2), three of four participants who did not located the 14th TA also chose to defend (Fisher's Exact P = .722). In other words, these groups were acting as though they were operating under worst case, rather than best case assumptions when dealing with uncertainty. The difference in course of action selections for group versus individual unaided cells that did not locate the 14th TA (three of four defending versus one of eight defending, respectively) is fairly significant (Fisher's Exact P = .069).

This result is especially surprising given the literature on group decision making. The standard finding is that group discussions and decision making tend to polarize individual decision making tendencies, i.e., make the groups more extreme than the-individuals (cf. Doise, 1969; Moscovici and Zavalloni, 1969). Therefore, if the individual tendency is to make best case assumptions regarding uncertainty and launch potentially risky attacks when there is the possibility of meeting an overwhelming force, one would expect that groups would behave even more extremely in this direction not less. In other words, it would be expected that the group's tendency to attack would be even greater than that of the individuals. The present data lie in the opposite direction and run counter to the well-established literature on group decision making.

Further, unlike the individual participants' results, in the group unaided condition, whether or not the groups have resolved uncertainty has no effect on course of action selection. One possible explanation for this is that groups build their own consensus regarding uncertainty and treat those conclusions as if they are true.

Interestingly, the group/unaided condition's tendency toward defense compared to the individual/unaided condition is in the direction of the experimental hypothesis, i.e., that group decision makers will make better decisions than individual unaided ones. However, given the pattern of results, it is difficult to use these results to claim support for the hypothesis.

An examination of the group/aided condition shows both a similar and a shockingly different pattern to the group/unaided condition. Like the group/unaided condition, whether or not the participants have resolved uncertainty has no effect on course of action selection. However, unlike the group/unaided condition (or either of the individual conditions), all group/aided cells chose to attack, regardless of whether or not they found the 14th TA.

Here, not only do participants in the group/aided condition appear to be best-casing unresolved uncertainty, they also appear to be best-casing "certainty." In other words, even when the 14th TA has been located in their sector, participants in this condition still act as though the situation favors them.

In this regard, the group/aided condition is different than both the individual/aided and group/unaided conditions. When the 14th TA is located, the group/aided condition participants attack, whereas the individual/aided participants defend. This difference is statistically significant (Fisher's Exact $P = .083$).

Similarly, the group/aided condition participants' overall tendency to attack is counter to the group/unaided condition participants' overall tendency to defend, regardless of whether participants have located the 14th TA. This difference is statistically significant (Fisher's Exact $P = .002$).

There is a possible explanation for what is going on in the group aided condition. Irving Janis discusses a phenomenon called "groupthink" (1972) which he argues affects groups of high-level decision makers. Among the effects of groupthink is the tendency for such decision makers to ignore uncertainty and view themselves as invulnerable. The results of the group/aided condition are consistent with this. Uncertainty had no effect on course of action chosen by participants in this cell (or for those in the group/unaided condition as well). Similarly, the fact that participants always chose to attack, regardless of whether the 14th TA was present is also consistent with this phenomenon.

The puzzling question is why participants in the group/unaided condition did not show these effects. While it is true that uncertainty had no effect on these participants, they tended to view themselves as "vulnerable" (i.e., chose to defend) even when it was uncertain that the 14th TA was present.

One explanation for this difference is that the wargaming aid itself can exacerbate the groupthink effect. Anecdotally, many participants in the group/aided condition would run the wargame with various elements of the enemy forces out of the picture. They justified this by arguing that those units would not participate in the defense against a particular friendly attack. In other words, these participants were using the aid to confirm their best-case assumptions and picking a course of action the aid showed to work.

This raises an important consideration to the aiding community. The present results suggest that groups can use aids to reinforce decision biases by manipulating the inputs to the aids. Therefore, it is imperative that aids not only support the decision maker's analysis, but also guard against potential biases that the aid might otherwise support. In other words, the aid must be "prescriptive" (helping the user make a good decision) as well as "personalized" (helping the user make the decision the way he wants to) (cf. Cohen et al., 1982).

It is interesting to note that this tendency to use an aid to bolster best-case assumptions was present only in the group condition and not the individual condition. It is not surprising that individuals would not be subject to "groupthink" effects, yet they could have just as easily used the aid to bolster best-case planning. The fact that groups did use the aid differently than individuals suggests that aid designers must take this variable into account when designing aids.

DISCUSSION

Results of Study 4 suggest that providing individual decision makers with partners, aids, or both produced no decrease in decision making time. While the original hypothesis called for such a decrease in time, the results suggest that such interventions can be made with no decrease but also no penalty in terms of time required to arrive at a decision (excluding the "aided-not used" data). On the other hand, providing such additional resources did result in an increase in the number of cells that located the 14th TA or resolved the uncertainties inherent in the situation. In other words, these cells developed a more accurate picture of the tactical situation. Also, as predicted, aids increased the number of courses of action considered by participants compared to those who received no aids.

The most important results in the present study concerned course of action selections. The first major result of interest was that at the individual decision making level, course of action selections were affected by whether or not participants resolved uncertainty (located the 14th TA). This is similar to Study 2. It made no difference whether or not participants had a wargaming aid at their disposal. As in Study 2, the tendency was that participants who did not locate the 14th TA appeared to be acting on best-case assumptions, namely choosing to attack in a potentially risky situation.

The results were reversed for the group decision makers. Here, whether or not participants located the 14th TA had no effect on course of action selection, but whether or not they used the wargaming aid did. Further, the group/unaided participants' tendency to defend suggested that they were adopting a worst-case treatment of uncertainty, contrary to the individual decision makers and group/aided decision makers, and contrary to what the literature on group decision making would predict. On the other hand, the group/aided participants' tendency to attack suggested that they were using the aid to bolster best-case assumptions.

At the individual level, it makes sense that both the aided and unaided participants would chose to defend when they locate the 14th TA in their sector. The aid clearly shows that an attack by the friendly forces would be defeated, whereas defending against an enemy attack would hold. Similarly, given the overwhelming size of the enemy threat, it is not difficult for unaided individuals to assess that attack is not feasible.

For participants who do not locate the 14th TA, the decision to attack would be justified under best-case reasoning. As studies 2 and 4 suggest, such reasoning did occur. What is particularly interesting is how individuals treated such uncertainty when provided with an aid. While the aid provided participants with the opportunity to perform "what if" reasoning, it seems clear that such reasoning occurs by manipulating variables the decision maker already knows about and not by including potential unforeseen variables. In other words, participants appeared to use the aid to manipulate inputs they had confirmed as being present in the scenario (e.g., confirmed friendly and enemy units and their strengths) rather than to explore the potential effects of variables that were unconfirmed but potentially important. It is interesting that no participant who did not locate the 14th TA ever wargamed a course of action that included the 14th TA as part of the enemy threat (this was true at the group level as well). This suggests that the participants may have

allocated their time disproportionately to wargaming scenarios that were biased in the favorable direction.

This finding is consistent with research on planners by Leddo and Govedich (1986). Leddo and Govedich found that planners typically give little consideration to unforeseen factors that can impact their plans. Also, in the Leddo and Govedich study, planners did not challenge the validity of assumptions they were making in planning. This is consistent with the present findings in that participants who were wargaming courses of action without the 14th TA never challenged their assumption that the 14th TA might actually commit into their sector.

This suggests that at the individual level, decision aids need to help decision makers challenge their assumptions, particularly their "best case" ones. One technique for doing this might be through the use of the "Conflict Resolution" technique (Leddo et al., 1990). The Conflict Resolution technique is a question and answer procedure that queries the decision maker on the reasons behind his choices. The goal is to ferret out assumptions. The decision maker is then asked to generate cases in which those assumptions might prove false, and what the impact on the decision would be when this occurs. This technique is straightforward enough so as to be implementable in an automated (or other) aid.

At the group level, it is somewhat baffling that uncertainty resolution had no effect on course of action selection. For the group/unaided condition, it makes sense that those who located the 14th TA would defend. It is perhaps also plausible that those who did not locate the 14th TA would also defend, but the fact they did while individuals did not is surprising, given the literature on group decision making cited earlier.

It does seem plausible that groups could develop their own consensus regarding the uncertain situation (which appeared to happen in both aided and unaided conditions). It is generally regarded in the social psychology literature that one of the functions of groups is to define "reality." If this is the case, then it could explain why course of action selection did not depend upon whether the 14th TA was located. Instead, the groups formed their own inference regarding the 14th TA and treated that inference as if it were certain.

What needs explaining is why the participants in the group/unaided conditions chose to defend while those in the aided condition chose to attack. The fact that these participants were conservative goes against traditional social psychology findings that groups tend amplify individual decision making tendencies (and the individual decision makers appeared to make best case rather than worst case treatments of uncertainty) and that groups of decision makers are often subject to overconfidence and hence take risks (the group-think effect). Further, given that the original mission was to attack, the group/unaided participants were selecting courses of action that went against this mission (although participants were instructed that they could pick the course of action they wanted).

Implications for Aiding

If the present results hold up in further experimentation, there are some powerful implications for the development and dissemination of aiding, as well as the structuring of individual and group decision making activities.

First, for unaided decision makers, it appears as though group decision makers are more conservative (in the Army tactical planning context) than individual decision makers. As it turns out, the optimal decision in the present study was to defend. Therefore, the group unaided decision makers came closest to making the "best" decisions. It would be interesting to have run the Bad to Good condition of Study 2 where the 14th TA did not enter the participants' sector. There, the optimal course of action would be to continue the attack. The question there would be whether group/unaided participants would be conservative and defend as well or would they choose to attack, even when they located the 14th TA outside their sector? If participants continued to be needlessly conservative in that situation, then it suggests that an Army reward structure that allowed for greater risk-taking might be appropriate.

For aided decision makers, it appears as though group decision makers are overly prone to making best case assumptions regarding their plans which leads to risky decision making. Here, the aid itself appears to be the culprit as decision makers use that aid to bolster their decision making. If this is so, it suggests that aids used in planning need to alert the user and help him consider cases where best case assumptions fail. Wargaming aids should not only allow for playing out scenarios that confirm expectations (i.e., exploring one branch of a decision tree) but should also provide for examining the impact of the major uncertainties with respect to the entire situation, thus providing the decision maker with a better understanding of the conditional probabilities of scenarios and associated outcomes as a function of assumptions adopted.

There is some evidence (cf. Tolcott, Marvin, and Bresnick, 1989), that aiding can reduce inference biases. Tolcott et al. found that educating people on inference biases, coupled with graphic displays that highlight uncertainty can result in an attenuation of the "confirmation bias" where people draw strong conclusions in the face of limited and even conflicting information. The Tolcott et al. results suggest that displays that highlighted the missing 14th TA (possibly similar to the graphic displays used in Study 3 that portrayed enemy strength using size of the symbol) and the effects that unit could have on course of action success might prove effective in attenuating reasoning by best case assumptions.

The results of the present study make an important statement regarding future research and development in the tactical decision making field. Specifically, it cannot be assumed that results obtained looking at individual decision makers can generalize to group decision making and vice versa. It also cannot be assumed that individuals will use decision aids in the same manner as will groups. The present study suggests that individuals and groups make decisions differently and use aids differently, and developers of aids will have to consider the individual/group issue.

REFERENCES

Cohen, M.S., Bromage, R.C., Chinnis, J.O., Jr., Payne, J.W., and Ulvila, J.W. (1982, July). A personalized and prescriptive attack planning decision aid (Technical Report 82-4). Falls Church, VA: Decision Science Consortium, Inc.

Cohen, M.S. Tolcott, M.A., and McIntyre, J. (1987, April). Display tech niques for pilot interactions with intelligent avionics: A cognitive approach (Technical Report 87-6). Falls Church, VA: Decision Science Consortium, Inc.

Cyert, Richard M., March, James G. (1963). A Behavorial Theory of the Firm, Englewood Cliffs, NJ, Prentice Hall.

Doise, W. (1969). Intergroup relations and polarization of individual and collective judgments. Journal of Personality and Social Psychology, 12, 136-143.

Einhorn, H.J. and Hogarth, R.M. (1984, June). Ambiguity and uncertainty in probabilistic inference. Chicago, Illinois: University of Chicago, Graduate School of Business.

Friedman M. and Savage L.J. (1948). The Utility Analysis of Choices Involving Risk. J. Political Economy, 56, 279-304.

Helson, H. (1948). Adaption Level as a Basis for Quantitative Theory of Frames of Reference. Psychological Review, V55, 297-313.

Janis, I.L. (1972). Victims of groupthink. Boston, MA: Houghton Mifflin.

Kahneman, D., Slovic, P. and Tversky, A. Eds. (1982). Judgment Under Uncer tainty: Heuristics and Biases, Cambridge University Press, Cambridge.

Kahneman, D., and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47(2), 263-291.

Langer, E.J. The illusion of control. In D. Kahneman, P. Slovic, and A. Tversky (Eds.) (1982). Judgment under uncertainty: Heuristics and biases. pp. 231-238. New York: Cambridge University Press.

Leddo, J., Chinnis, J.O., Jr., Cohen, M.S., and Marvin, F.F. (1987, January). The influence of uncertainty and time stress on decision making (Technical Report 87-1). Falls Church, VA: Decision Science Consortium, Inc.

Loewenstein, G. and Linville, P. (1986, March). Expectation formation and the timing of outcomes: A cognitive strategy for balancing the conflicting incentives for savoring success and avoiding disappointment. Chicago, Illinois: University of Chicago, Graduate School of Business.

Moscovici, S. and Zavalloni, M. (1969). The group as a polarizer of atti tudes. Journal of Personality and Social Psychology, 12, 124-135.

Seligman, M.E.P. (1975). Helplessness: On depression, development and death. San Francisco: W.H. Freeman.

Simon, H. (1957). Models of Man: Social and Rational. New York: Wiley.

Slovic, P., Fischhoff, B., and Lichtenstein, L. (1982) Facts versus fears: Understanding perceived risk. In D. Kahneman, P. Slovic, and A. Tversky (Eds.), Judgment under uncertainty: Heuristics and biases. New York: Cambridge University Press, pp. 463- 489.

Tolcott, M.A., Marvin, F.F., and Bresnick, T.A. (1989 August). Situation assessment and hypothesis testing in an evolving situation. (Technical Report 89-3). Reston, VA: Decision Science Consortium, Inc.

Tversky, A. and Kahneman, D. (1981). The framing of decisions and the rationality of choice. Science, 211, 453-458.

Tversky, A. and Kahneman, D. (1974). Judgment Under Uncertainty: Heuristics and Biases. Science, 1895, 1124-1131.

Vane, R.R., Black, P.K., Laskey, K.B., and Bresnick, T.A. (1990, May). Force allocation and comparison tool (FACT) (Technical Report 90-5). Reston, VA: Decision Science Consortium, Inc.

White, R.W. (1959) Motivation reconsidered: The concept of competence. Psychological Review, 66, 297-333.

APPENDIX A

AIDING SUGGESTIONS BY SUBJECTS

APPENDIX A

AIDING SUGGESTIONS BY SUBJECTS

SUBJECT	CONDITION	AID 1	AID 2	AID 3	AID 4	AID 5	AID 6
3	NSBG	Event Template	Situation Assessment	War Gaming Aid			
5	NSBG	Event Matrix	Critical Event	Execution Matrix	Synch Matrix	Decision Matrix	
8	NSBG	Task Organizer	Situation Template	Evt Hist/Battle Flow			
17	NSBG	Fire Support Plan	CAS Sorties Avail	Nuke/Caslty Analysis	Terrain Analysis	Info Screening Aid	
19	NSBG	Graphic Intelligence					
24	NSBG						
27	NSBG	Summary of Estimates					
31	NSBG	Enemy Identifier	Vehicle Concentratn	Link with Scouts			
15	NSGB	Decision Support Mtx	Staff Briefings				
16	NSGB	Data Link with G2/G3	Automated Database				
20	NSGB	Force Ratio Analysis	Automated Intel	Graphic Force Disply			
21	NSGB	Information Summary	Automated Displays				
22	NSGB	Situation Assessment	G1/G4 Status	Div/Corps Interchng	OB		
25	NSGB	Situation Template	Doctrinal Template	Personality Profiles	Range Estimators	Equip Capability	Equip Status Charts
29	NSGB	Situation Updater					
33	NSGB	Unit ID	Intel Estimate	Eng Assessment	Terrain Analysis	Logistics Analysis	
1	TSBG	Info Relevance	Situation Template	War Gaming Aid			
2	TSBG	Doctrinal Template	Terrain Analysis	Logistics Aid	COA Analyzer		
9	TSBG	Visual Status Board	Combat Power Anlys	Templating			
11	TSBG	COA Overlay	Friendly Equip Plot	Enemy Task Org	G-2 Brief	Message Plot	German Language Help
14	TSBG	Plot Intel By Area	Wargame Enemy COA	ADFATDS	G4 Analysis Fr COA		
23	TSBG	Intel on Enemy Intnt					
28	TSBG	Situation Updater	Info Analyzer				
32	TSBG	Status Charts	Visual Cht of En Str				
4	TSGB	Task Org of Units	Enemy OB	Mission Chart	G-3 Milestone Chart	COA Sketches	War Gaming
7	TSGB	Situation Assessment	Info Condenser				
10	TSGB	Info Condenser	Situation Updater				
12	TSGB	Updated Messages	Intel Plot By Area	Staff Recmndtns COA			
18	TSGB	Info Condenser	Flow of Battle Aid	War Gaming			
26	TSGB	Doctrinal Template	G2/G3 Aid	IPB Overlays			
30	TSGB	Checklists	Automated Database	Timeline of Tasks	Simplified Intel	Terrain Info	Decision Template
34	TSGB	Color Code Status	Previous En Action	CDR's Mission	Sorties	IPB Products	Enemy Update Aid